现代生物技术前沿

夏其昌 曾嵘 等 编著

蛋白质化学与蛋白质组学

科学出版社
www.sciencep.com

Digitized by the Internet Archive
in 2011 with funding from
Institute of Botany, CAS and Internet Archive

http://www.archive.org/details/danbaizhihuaxuey00xiaq

中科院植物所图书馆
S0003766
58.17421
468

现代生物技术前沿

夏其昌 曾嵘 等 编著

蛋白质化学与蛋白质组学

281.23

科学出版社
北京

内 容 简 介

本书系统论述了蛋白质化学基础理论和实验技巧，也反映了蛋白质组学研究的最新成果。内容包括：蛋白质的表征，蛋白质的组成分析和序列测定，与此相关的实验方法，包括各种色谱、电泳、质谱技术等，以及应用在蛋白质表征研究和基因工程产品的质检方面的实际范例。在蛋白质组学领域介绍了基本概念、样品制备、双向凝胶电泳的图像分析和定量分析、质谱等常规方法，并介绍了国际上最新的多维技术在研究中的应用；同时充分体现了生物信息学在蛋白质组研究中的重要性。

本书可作为生物学、医学、化学专业大学生，研究生和教学人员的参考书，也是从事生物化学、分子生物学、医学等领域中分离分析工作人员的参考书。

图书在版编目(CIP)数据

蛋白质化学与蛋白质组学/夏其昌，曾嵘等编著. —北京：科学出版社，2004.4

(现代生物技术前沿)

ISBN 7-03-012401-4

Ⅰ.蛋… Ⅱ.①夏…②曾… Ⅲ.蛋白质-研究 Ⅳ.Q51

中国版本图书馆 CIP 数据核字(2003)第 100086 号

责任编辑：莫结胜　杨兆弘　谢灵玲／责任校对：宋玲玲

责任印制：安春生／封面设计：王　浩　陈　敬

科学出版社 出版

北京东黄城根北街16号

邮政编码：100717

http://www.sciencep.com

源海印刷有限责任公司印刷

科学出版社发行　各地新华书店经销

*

2004 年 4 月第 一 版　开本：787×1092　1/16

2004 年 4 月第一次印刷　印张：36 1/4

印数：1—4 000　字数：828 000

定价：75.00 元

(如有印装质量问题，我社负责调换〈新欣〉)

《蛋白质化学与蛋白质组学》编写人员名录

（按姓氏音序排序）

曹兴军　中国科学院上海生命科学研究院生物化学与细胞生物学研究所
蛋白质组研究分析中心　上海 200031

车发云　Department of Molecular Pharmacology, Albert Einstein College of Medicine, New York, NY, USA

陈益强　中国科学院计算技术研究所　北京 100080

丁士健　中国科学院上海生命科学研究院生物化学与细胞生物学研究所
蛋白质组研究分析中心　上海 200031

付　岩　中国科学院计算技术研究所　北京 100080

龚蓁蓁　中国科学院上海生命科学研究院生物化学与细胞生物学研究所
上海 200031

江晓胜　中国科学院上海生命科学研究院生物化学与细胞生物学研究所
蛋白质组研究分析中心　上海 200031

靳文海　中国科学院大连化学物理研究所　国家色谱中心　大连 116023

李　辰　中国科学院上海生命科学研究院生物化学与细胞生物学研究所
蛋白质组研究分析中心　上海 200031

李荣霞　中国科学院上海生命科学研究院生物化学与细胞生物学研究所
蛋白质组研究分析中心　上海 200031

李素君　中国科学院上海生命科学研究院生物信息中心　上海 200031

李亦学　中国科学院上海生命科学研究院生物信息中心　上海 200031

刘　韬　Howard Hughes Medical Institute, University of Washington, Seattle, WA, USA

马丹军　中国科学院上海生命科学研究院生物化学与细胞生物学研究所
蛋白质组研究分析中心　上海 200031

阮宏强　中国科学院上海生命科学研究院生物化学与细胞生物学研究所
蛋白质组研究分析中心　上海 200031

邵晓霞　中国科学院上海生命科学研究院生物化学与细胞生物学研究所
蛋白质组研究分析中心　上海 200031

沈海芳　中国科学院上海生命科学研究院药物研究所　上海 201203

史　律　中国科学院上海生命科学研究院生物化学与细胞生物学研究所
蛋白质组研究分析中心　上海 200031

宋金芳　中国科学院上海生命科学研究院生物化学与细胞生物学研究所
蛋白质组研究分析中心　上海 200031

唐柳娅　中国科学院上海生命科学研究院生物化学与细胞生物学研究所
　　　　蛋白质组研究分析中心　　上海 200031
屠成剑　中国科学院上海生命科学研究院生物化学与细胞生物学研究所
　　　　蛋白质组研究分析中心　　上海 200031
屠红旻　Biocenter and Department of Medical Biochemistry & Molecular
　　　　Biology, University of Oulu, Oulu, Finland
吴家睿　中国科学院上海生命科学研究院生物化学与细胞生物学研究所
　　　　上海 200031
夏其昌　中国科学院上海生命科学研究院生物化学与细胞生物学研究所
　　　　蛋白质组研究分析中心　　上海 200031
徐来根　中国科学院上海生命科学研究院生物化学与细胞生物学研究所
　　　　蛋白质组研究分析中心　　上海 200031
薛卫华　中国科学院上海生命科学研究院生物化学与细胞生物学研究所
　　　　蛋白质组研究分析中心　　上海 200031
杨一鸣　中国科学院上海生命科学研究院药物研究所　　上海 201203
俞利荣　Laboratory of Proteomics and Analytical Technologies SAIC-Frederick,
　　　　Inc. Natinal Cancer Institute at Frederick, MD, USA
曾　嵘　中国科学院上海生命科学研究院生物化学与细胞生物学研究所
　　　　蛋白质组研究分析中心　　上海 200031
张　雷　中国科学院上海生命科学研究院生物化学与细胞生物学研究所
　　　　蛋白质组研究分析中心　　上海 200031
张丽华　Medical Information Technology Inc., Boston, USA
周　虎　中国科学院上海生命科学研究院生物化学与细胞生物学研究所
　　　　蛋白质组研究分析中心　　上海 200031
邹汉法　中国科学院大连化学物理研究所　国家色谱中心　大连 116023

前　　言

我曾在 1997 年 10 月出版了《蛋白质化学研究技术与进展》一书。此书后来又重印了一次，这反映有一定的需求，但目前已完成其历史任务。因为，这几年中，蛋白质领域发生了翻天覆地的变化，进入了蛋白质组学的时代。在中国科学院，上海生命科学研究院和生物化学与细胞生物学研究所于 2000 年共建了蛋白质组研究分析中心。近年来，更多的研究单位和大专院校建立了蛋白质组学中心或蛋白质组学重点实验室。

蛋白质化学的研究是从细胞中成千上万种蛋白质中分离纯化到一种蛋白质来研究它；而在蛋白质组学研究中，是欲将细胞中全部的蛋白质系统地分离和鉴定，在整体水平上研究其结构和功能。有人说，早期蛋白质研究是个体作坊式的，而目前的蛋白质组研究是在团队水平上进行。由于工作量之大和高通量、高难度，不是少数几个人所能胜任的，这也要求研究者是团队式的。

30 多年前，很多医学家已经考虑到疾病不是单因素而是多种因素综合作用的结果，他们发展了现代双向电泳技术来研究这些复杂的因素，看到了可喜的前景。当时犹如看天上众多的星星，但理不出头绪。后来发展的电泳图像分析和质谱技术，才解决了这些疑难问题。

我近年常用"从蛋白质化学到蛋白质组学"这一题目讲课，这也反映一种历程。国际上一些科学家用疑惑的眼光询问，也有很多人表示赞同。事实上，不少蛋白质组学专家，过去也是蛋白质化学家。当然，也有不少蛋白质组学专家来自医学界。

1997 年，我应邀参加在澳大利亚举行的 Lorne 蛋白质结构和功能讨论会，当时 Williams 的报告"The Establishment Phase for the Australian Proteome Analysis Facility"(即著名的 APAF)使我很受启发，次年我访问美国一些实验室时，特别注意有关内容，并请顺访上海的蛋白质化学专家潘裕敬博士作了蛋白质组的报告，重点在双向电泳。当时，我们已有 CE、Mimi-IEF、FFE 等多种电泳技术，但没有建立双向电泳的要求，也不太明白为何在国际有名的电泳杂志上总有一半论文涉及 2-DE。当对蛋白质组学有所了解后，也就明白了。

金善炜教授多次建议我：做蛋白质研究一定要有质谱！在后来工作中，还得到质谱专家杨一鸣进一步的指点和帮助。

我步入蛋白质化学这一领域首先是在复旦大学的课堂上听邹承鲁、钮经义先生和曹天钦先生的精彩讲课，当时很多名教授为大学生讲系列的基础课。进入中国科学院上海生物化学研究所后，得益于王应睐先生和戚德芳先生的直接教诲。值得指出的是：在著名的生化所这样的研究氛围中，我受到国内一批享有盛名的研究蛋白质的前辈熏陶；同时也获益于同事们有益的交流讨论，他们在当时生活和研究条件极其艰难的情况下开展工作。

蛋白质组学这一新的热点学科在国内外都有不同程度的理解。可以作为大学科来组织，也可以作为一种研究方法在小范围内对所感兴趣的课题进行研究。但有一点是肯定的，即大家都对蛋白质组学寄予厚望，并在不同水平和不同范畴上进行工作，这是值得庆

贺的。如今已有全球性的蛋白质组机构——人类蛋白质组组织(Human Proteome Organization, HUPO),蛋白质组学的专业性刊物也不断涌现,一派欣欣向荣景象。

我们在《蛋白质化学研究技术与进展》的基础上,结合当前的研究进展,写作了本书,如果本书能在我国蛋白质组学的发展中,引起人们更多的兴趣,起到普及和引导入门的作用,我们将感到非常欣慰。

夏其昌
2003年4月21日
中国科学院　上海生命科学研究院
生物化学与细胞生物学研究所
蛋白质组研究分析中心
http://www.proteomics.ac.cn

目　录

前言

上篇　蛋白质化学

第一章　蛋白质的表征 …… 3
第一节　蛋白质结构的基本概念 …… 3
第二节　蛋白质的纯度 …… 6
第三节　蛋白质的定量 …… 7
一、光吸收法 …… 8
二、双缩脲法 …… 8
三、福林-酚法 …… 8
四、考马斯亮蓝 G-250 法 …… 9
五、氨基酸分析法 …… 9
第四节　蛋白质的脱盐(除去盐和非共价结合的小分子) …… 9
第五节　蛋白质的分子质量测定 …… 10
第六节　蛋白质的等电点测定 …… 11
第七节　氨基酸定量分析 …… 12
第八节　肽　谱 …… 14
第九节　质　谱 …… 16
第十节　蛋白质及多肽的序列测定和末端分析 …… 18
一、N 端分析 …… 19
二、C 端的测定 …… 19
第十一节　蛋白质的二硫键分析 …… 19
一、不同同位素标记法 …… 20
二、4-乙烯吡啶标记法 …… 20
三、直接色谱法 …… 21
四、质谱法 …… 21
第十二节　突变点的分析 …… 22
第十三节　原位分析和快速分析 …… 24
一、原位分析 …… 24
二、快速分析 …… 24
第十四节　蛋白质翻译后修饰的分析 …… 25
参考文献 …… 25
第二章　蛋白质结构分析 …… 27
第一节　蛋白质样品的准备 …… 27

一、SDS-PAGE 纯化的样品 …… 27
二、HPLC 纯化的样品 …… 27
三、蛋白质样品中的二硫键和巯基(—SH) …… 27
第二节　蛋白质的化学裂解 …… 28
一、裂解于甲硫氨酸（Met） …… 28
二、裂解于半胱氨酸(Cys) …… 29
三、裂解于色氨酸（Trp） …… 30
四、部分酸水解 …… 30
五、选择性及限制性化学裂解 …… 30
第三节　蛋白水解酶酶解 …… 31
一、常用蛋白水解酶 …… 31
二、酶解操作条件 …… 31
三、讨论 …… 32
第四节　蛋白质中存在着封闭的 N 端 …… 33
一、除去蛋白质中封闭的 *N*-乙酰丝氨酸和 *N*-乙酰苏氨酸 …… 33
二、除去 *N*-甲酰甲硫氨酸 …… 33
三、除去 N 端的焦谷氨酸残基（pyroglutamate） …… 33
四、除去其他的 *N*-乙酰氨基酸 …… 34
五、选择性纯化 N 端封闭的肽 …… 34
第五节　蛋白质中的 C 端残基的测定 …… 34
一、羧肽酶法 …… 34
二、化学降解法 …… 35
第六节　蛋白质中一些特殊肽段的检出 …… 35
一、含巯基肽的分离纯化和分析 …… 35
二、含芳香族氨基酸肽的分离纯化和鉴定 …… 37
三、含 Lys-肽的分离纯化和鉴定 …… 40
参考文献 …… 42
附录一　微量蛋白质实验室注意事项 …… 44
附录二　常用试剂处理 …… 44
第三章　氨基酸组成分析 …… 46
第一节　氨基酸组成分析的目的 …… 46
第二节　蛋白质的水解方法 …… 46
第三节　特殊氨基酸的保护和定量 …… 48
第四节　氨基酸组成原位分析 …… 50
第五节　衍生方法及原理 …… 50
第六节　测定氨基酸组成的实验步骤 …… 52
一、OPA 法 …… 52
二、PTC-AA 法 …… 53
三、茚三酮法 …… 54

四、氨基酸直接分析法 …… 55
第七节　结　语 …… 55
参考文献 …… 56
第四章　蛋白质和多肽的氨基酸序列测定 …… 57
第一节　N端分析原理 …… 57
一、耦联 …… 57
二、裂解 …… 57
三、转化 …… 58
四、PTH-氨基酸的鉴定 …… 58
第二节　N端蛋白质序列仪 …… 58
一、液相旋转杯序列仪 …… 58
二、固相序列分析仪 …… 59
三、气相序列仪 …… 59
第三节　影响N端Edman反应裂解率的因素 …… 60
第四节　测序前样品处理 …… 62
一、纯度鉴定 …… 62
二、脱盐 …… 62
三、巯基修饰方法 …… 62
四、去除N端封闭的基团的方法 …… 63
第五节　手工N端测序 …… 64
第六节　C端序列分析 …… 67
一、C端分析原理 …… 67
二、C端蛋白质序列仪 …… 69
三、C端测序的样品前处理 …… 69
四、影响C端测序反应产率的因素 …… 72
五、结语和展望 …… 72
参考文献 …… 73
第五章　薄层等电聚焦 …… 74
第一节　原理 …… 74
第二节　薄层IEF …… 75
一、仪器 …… 75
二、试剂和储存液配制 …… 75
三、制胶模具的准备 …… 75
四、凝胶的制备 …… 76
五、加样 …… 76
六、电泳 …… 76
七、染色和脱色 …… 76
八、凝胶保存 …… 77
第三节　等电聚焦中的注意事项 …… 77

第四节 IEF 的优缺点 …… 77
第五节 实例 …… 78
参考文献 …… 79
第六章 SDS-聚丙烯酰胺凝胶电泳 …… 80
第一节 概况 …… 80
第二节 材料 …… 80
一、设备和材料 …… 80
二、试剂 …… 81
第三节 基本操作步骤 …… 81
一、聚胶前准备 …… 81
二、不连续系统下层分离胶的聚合 …… 81
三、不连续系统中的上层浓缩胶和连续系统凝胶的聚合 …… 82
四、加样 …… 82
五、电泳 …… 82
六、染色和脱色 …… 82
七、干胶 …… 84
第四节 几种不同凝胶电泳系统及其应用 …… 84
一、Leammli 系统 …… 84
二、大孔胶系统 …… 86
三、Tris-Tricine 系统 …… 87
第五节 注意事项 …… 88
一、形成凝胶的试剂要有足够的纯度 …… 88
二、激活剂的用量 …… 89
三、温度的影响 …… 90
四、氧气的影响 …… 90
五、凝胶添加剂 …… 90
六、聚胶时间 …… 90
七、单体的浓度%T,%C …… 90
八、操作注意点 …… 90
参考文献 …… 91
第七章 液相电泳 …… 92
第一节 毛细管电泳 …… 92
一、概况 …… 92
二、材料 …… 95
三、基本操作步骤 …… 96
四、毛细管电泳分离模式及应用 …… 97
五、注意事项 …… 104
六、结束语 …… 104
第二节 连续自由流电泳 …… 105

一、概况 …… 105
二、材料 …… 109
三、自由流电泳操作步骤 …… 110
四、分离模式 …… 111
五、自由流电泳的应用 …… 113
六、自由流电泳在微重力下的特点 …… 115
七、展望 …… 116
第三节 液相等电聚焦 …… 117
一、旋转式等电聚焦 …… 117
二、循环式等电聚焦 …… 119
三、制备型连续流等电聚焦 …… 122
参考文献 …… 123
第八章 高效液相色谱 …… 127
第一节 高效液相色谱的类型及分离原理 …… 127
第二节 高效液相色谱的装置 …… 131
一、输液系统 …… 131
二、进样系统 …… 132
三、分离系统 …… 132
四、检测系统 …… 136
第三节 各类高效液相色谱应用实例 …… 137
一、反相高效液相色谱 …… 137
二、高效离子交换色谱 …… 138
三、高效疏水作用色谱 …… 138
四、高效排阻色谱 …… 138
五、高效亲和色谱 …… 139
第四节 讨论 …… 140
参考文献 …… 142
第九章 质谱 …… 144
第一节 引言 …… 144
第二节 四极和离子阱质谱 …… 144
一、四极质谱仪(quadrupole mass spectrometry) …… 144
二、四极离子阱质谱(quadrupole ion trap mass spectrometry) …… 147
第三节 飞行时间质谱 …… 154
第四节 傅里叶转换回旋共振质谱 …… 160
一、基本原理 …… 160
二、质量分辨率、质量精度和质量范围 …… 163
参考文献 …… 165
第十章 毛细管电泳-质谱联用 …… 167
第一节 介绍 …… 167

一、概况 …… 167
二、CE/MS联用的接口 …… 169
三、其他联用方式 …… 172
四、CE/MS的限制性因素及其改进方法 …… 173
五、应用 …… 177
第二节　材料 …… 179
一、试剂和溶剂 …… 179
二、多肽和蛋白质 …… 180
三、酶 …… 180
四、仪器 …… 181
第三节　方法 …… 181
一、胰蛋白酶酶解 …… 181
二、血红蛋白的脱盐和纯化 …… 181
三、蛋白磷酸酶去磷酸化 …… 181
四、蛋白质二硫键的羧酰胺甲基化及产物的HPLC纯化和蛋白酶酶解 …… 182
五、*N*-糖苷酶F酶解 …… 182
六、α-甘露糖苷酶酶解 …… 182
七、正离子非共价涂层毛细管电泳 …… 182
八、毛细管电泳-质谱联用 …… 182
九、蛋白质序列数据库搜索 …… 183
十、特殊肽段的选择性离子监测 …… 184
十一、毛细管等电聚焦-质谱联用 …… 184
第四节　结果 …… 185
一、多肽和蛋白质的分析与鉴定 …… 185
二、蛋白质点突变的鉴定 …… 190
三、蛋白质磷酸化的鉴定 …… 193
四、蛋白质*N*-糖基化和氧化的鉴定 …… 194
五、蛋白质混合物的毛细管等电聚焦-质谱联用分析 …… 199
第五节　展望 …… 203
参考文献 …… 203
第十一章　毛细管电泳在蛋白质磷酸化和糖基化分析中的应用 …… 208
第一节　毛细管电泳分析多肽和蛋白质磷酸化 …… 208
一、简介 …… 208
二、材料 …… 209
三、方法 …… 209
四、结果和讨论 …… 210
第二节　毛细管电泳分析单糖和寡糖 …… 212
一、简介 …… 212
二、材料 …… 216

三、方法 …………………………………………………………………………… 216
四、结果和讨论……………………………………………………………………… 217
第三节　毛细管电泳分析糖蛋白微不均一性 ……………………………………… 221
一、简介 …………………………………………………………………………… 221
二、材料 …………………………………………………………………………… 223
三、方法 …………………………………………………………………………… 223
四、结果和讨论……………………………………………………………………… 223
第四节　结语和展望………………………………………………………………… 225
参考文献 ……………………………………………………………………………… 226

下篇　蛋白质组学

第十二章　导论………………………………………………………………………… 233
一、蛋白质组学研究的历史和背景 ……………………………………………… 233
二、蛋白质组学的特点与难点 …………………………………………………… 234
三、蛋白质组学发展趋势 ………………………………………………………… 239
参考文献 ……………………………………………………………………………… 240
第十三章　样品的全息制备………………………………………………………… 242
第一节　双向凝胶电泳常规样品制备及其改进 …………………………………… 242
一、概述 …………………………………………………………………………… 242
二、样品裂解液中各种成分分析…………………………………………………… 243
三、双向凝胶电泳常规样品制备…………………………………………………… 245
四、蛋白质顺序提取法 …………………………………………………………… 245
五、注意事项 ……………………………………………………………………… 246
第二节　培养细胞蛋白质样品的制备……………………………………………… 246
一、细胞破碎的处理方法 ………………………………………………………… 246
二、培养细胞总蛋白质的提取 …………………………………………………… 247
第三节　组织样品制备 ……………………………………………………………… 249
一、概述 …………………………………………………………………………… 249
二、动物组织样品的制备 ………………………………………………………… 249
三、植物组织样品的制备 ………………………………………………………… 260
四、注意事项 ……………………………………………………………………… 261
第四节　分泌蛋白质样品的制备 …………………………………………………… 261
一、真核细胞的分泌蛋白质样品制备……………………………………………… 262
二、细菌分泌蛋白质样品制备 …………………………………………………… 263
第五节　体液蛋白质样品的制备 …………………………………………………… 263
一、血清蛋白质样品的制备 ……………………………………………………… 263
二、脑脊液蛋白质样品的制备 …………………………………………………… 264
三、总结 …………………………………………………………………………… 265
参考文献 ……………………………………………………………………………… 265

第十四章　双向凝胶电泳…………………………………………………………………… 269
第一节　介绍 ………………………………………………………………………… 269
一、双向凝胶电泳简介 ………………………………………………………… 269
二、IPG-DALT 双向凝胶电泳…………………………………………………… 270
第二节　试剂、设备和溶液配制 ………………………………………………… 275
一、试剂 …………………………………………………………………………… 275
二、仪器 …………………………………………………………………………… 275
三、溶液配制 ……………………………………………………………………… 276
第三节　实验操作…………………………………………………………………… 277
一、样品制备 ……………………………………………………………………… 277
二、蛋白质定量…………………………………………………………………… 278
三、双向凝胶电泳 ………………………………………………………………… 279
第四节　胶上蛋白质的检测 ……………………………………………………… 280
一、介绍 …………………………………………………………………………… 280
二、染色方法 ……………………………………………………………………… 283
参考文献 ……………………………………………………………………………… 286
第十五章　电泳图谱的图像分析………………………………………………………… 290
第一节　介绍 ………………………………………………………………………… 290
第二节　PDQuest 软件简介 ……………………………………………………… 290
第三节　PDQuest 软件的使用 …………………………………………………… 291
一、材料 …………………………………………………………………………… 291
二、方法 …………………………………………………………………………… 292
第四节　双向凝胶电泳软件的实际应用 ………………………………………… 299
一、双向凝胶电泳的重复性和定量定性差异分析………………………………… 299
二、双向凝胶电泳矢量图 ………………………………………………………… 304
第五节　双向凝胶电泳数据库和网上比较 ……………………………………… 305
一、双向凝胶电泳网络数据库资源……………………………………………… 305
二、双向凝胶电泳数据库的构建………………………………………………… 308
三、双向电泳凝胶的网上比较 …………………………………………………… 309
参考文献 ……………………………………………………………………………… 310
第十六章　生物质谱技术和蛋白质鉴定………………………………………………… 311
第一节　生物质谱基本原理和工作模式 ………………………………………… 311
第二节　质谱法分析完整蛋白质和多肽 ………………………………………… 316
第三节　质谱法对蛋白质和多肽一级结构的分析及鉴定 ……………………… 317
一、质谱分析多肽序列的原理 …………………………………………………… 317
二、质谱法结合数据库检索对多肽和蛋白质进行鉴定 ………………………… 318
三、多肽从头测序………………………………………………………………… 325
第四节　质谱前蛋白质或多肽样品制备方法和关键步骤 ……………………… 328
参考文献 ……………………………………………………………………………… 335

第十七章　蛋白质组研究中的定量方法…… 340
第一节　介绍 …… 340
一、概述 …… 340
二、化学标记定量方法 …… 341
第二节　材料 …… 346
一、试剂和溶剂 …… 346
二、多肽、蛋白质和细胞 …… 347
三、仪器 …… 347
第三节　方法 …… 347
第四节　结果和讨论 …… 349
第五节　展望 …… 356
参考文献 …… 356
第十八章　蛋白质组研究中的翻译后修饰分析…… 358
第一节　磷酸化蛋白质组研究 …… 358
一、介绍 …… 358
二、材料 …… 369
三、方法 …… 370
四、结果 …… 371
五、展望 …… 372
第二节　糖基化蛋白质组研究 …… 374
一、介绍 …… 374
二、材料 …… 383
三、方法 …… 383
四、结果和讨论 …… 384
五、展望 …… 388
参考文献 …… 389
第十九章　亚细胞蛋白质组学…… 394
第一节　概论 …… 394
一、亚细胞蛋白质组学内涵 …… 394
二、亚细胞蛋白质组学的意义 …… 395
三、亚细胞蛋白质组学的技术基础 …… 396
四、亚细胞蛋白质组学研究进展 …… 404
五、总结 …… 405
第二节　细胞膜的蛋白质组学研究进展 …… 406
一、细胞膜的结构与功能 …… 406
二、膜蛋白质的特点 …… 406
三、膜蛋白质的双向凝胶电泳技术研究 …… 407
四、膜蛋白质的非双向凝胶电泳技术研究 …… 408
五、总结 …… 408

第三节　高尔基体蛋白质组学研究进展 …… 408
一、高尔基体的结构和功能 …… 408
二、高尔基体的蛋白质组学研究 …… 409
三、蛋白质组学研究发现新的高尔基体蛋白质 …… 410
四、不同功能状态高尔基体的比较蛋白质组学研究 …… 410
五、总结 …… 411
第四节　核孔复合体蛋白质组学研究进展 …… 411
一、核孔复合体的结构和功能 …… 411
二、酵母核孔复合体蛋白质组学研究 …… 412
三、哺乳动物核孔复合体蛋白质组学研究 …… 413
四、总结 …… 413
第五节　核仁蛋白质组学研究进展 …… 414
一、核仁的结构与功能 …… 414
二、核仁的蛋白质组学研究进展 …… 415
三、总结 …… 415
第六节　线粒体蛋白质组学研究 …… 415
一、线粒体蛋白质组学研究进展 …… 416
二、材料 …… 420
三、方法 …… 421
四、结果 …… 422
第七节　展望 …… 422
参考文献 …… 424
第二十章　蛋白质组研究中的非凝胶技术 …… 432
第一节　概况 …… 432
第二节　非凝胶色谱技术 …… 433
一、一维色谱及其与质谱联用技术 …… 433
二、多维色谱及其与质谱联用技术 …… 435
第三节　非凝胶电泳 …… 439
一、液相等电聚焦 …… 439
二、毛细管电泳 …… 444
第四节　毛细管电色谱及其在蛋白质组学研究中的初步应用 …… 444
一、毛细管电色谱简介 …… 444
二、毛细管电色谱在蛋白质和多肽分离中的应用 …… 448
第五节　展望 …… 449
参考文献 …… 450
第二十一章　蛋白质相互作用和蛋白质芯片 …… 456
第一节　蛋白质相互作用 …… 456
一、概况 …… 456
二、蛋白质相互作用研究策略和方法 …… 457

三、结语 …… 467
第二节 蛋白质芯片 …… 467
一、蛋白质芯片的分类 …… 468
二、蛋白质芯片的检测方法 …… 471
三、蛋白质芯片技术的应用及发展前景 …… 472
参考文献 …… 473
第二十二章 蛋白质组生物信息学 …… 477
一、大规模蛋白质组学实验信息的产生 …… 477
二、将实验信息处理成具有生物学意义的实验结果 …… 477
三、对实验结果的分析 …… 491
四、实验室信息管理系统 …… 506
五、总结 …… 507
参考文献 …… 507
第二十三章 蛋白质组学在各领域的应用 …… 511
第一节 蛋白质组学在微生物研究中的应用 …… 511
一、概述 …… 511
二、蛋白质组学在模式微生物研究中的应用 …… 512
三、蛋白质组学在病原微生物研究中的应用 …… 515
四、展望 …… 521
第二节 蛋白质组学在肿瘤研究中的应用 …… 521
第三节 蛋白质组学在神经系统研究中的应用 …… 528
一、介绍 …… 528
二、应用举例 …… 529
第四节 蛋白质组学在体液研究中的应用 …… 534
一、介绍 …… 534
二、体液的蛋白质组学研究现状 …… 536
三、展望 …… 544
第五节 蛋白质组学在其他领域的应用 …… 545
一、蛋白质组学在植物研究中的应用 …… 545
二、蛋白质组学在药物开发方面的研究 …… 547
三、结语 …… 548
参考文献 …… 548
后记 …… 559

上 篇

蛋 白 质 化 学

第一章　蛋白质的表征

第一节　蛋白质结构的基本概念

蛋白质是一类重要的生物高分子，英文为"Protein"，该词从希腊文转化而来，意思是"最原始的"，可见当时人们对蛋白质的认识已有相当的基础。

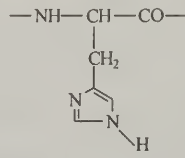
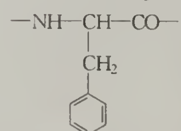
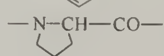
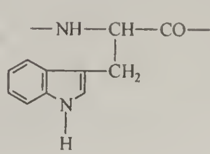
蛋白质有成千上万种，不同的蛋白质具有不同的生物功能。蛋白质是 20 种 α-氨基酸通过酰胺键（肽键）联成的长链分子，这种长链即所谓的肽链。许多蛋白质还包含有非肽链结构的其他组成成分，这种成分称为配基或辅基。

作为蛋白质组成成分的氨基酸一共有 20 种，它们的名称和结构列在表 1.1 中。为了表达蛋白质和肽结构的需要，每个氨基酸都有相应的符号表示，常用符号通常由氨基酸英文名的三个字母组成。一个蛋白质由上百个甚至上千个氨基酸组成，为了表达的方便，又设计出单字符号。

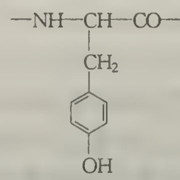
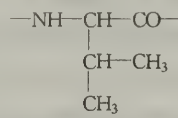
表 1.1　存在于蛋白质中的 20 种常见氨基酸

氨基酸	英文名	三字母符号	单字母符号	结　构	残基分子质量/Da	等电点
丙氨酸	alanine	Ala	A	—NH—CH—CO— | CH_3	71.08	6.00
精氨酸	argnine	Arg	R	—NH—CH—CO— | $(CH_2)_3$ | NH | NH_2—C=NH	156.19	10.76
天冬酰胺	asparagine	Asn	N	—NH—CH—CO— | CH_2 | $CONH_2$	114.10	
天冬氨酸	aspartic acid	Asp	D	—NH—CH—CO— | CH_2 | COOH	115.09	2.77
天冬酰胺或天冬氨酸		Asx	B			
半胱氨酸	cysteine	Cys	C	—NH—CH—CO— | CH_2SH	103.14	5.07
谷氨酸	glutamic acid	Glu	E	—NH—CH—CO— | CH_2 | CH_2—COOH	129.12	3.22

续表

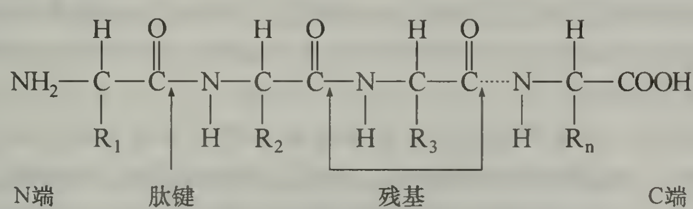
氨基酸	英文名	三字母符号	单字母符号	结　构	残基分子质量/Da	等电点
谷氨酰胺	glutamine	Gln	Q	—NH—CH—CO— CH_2 CH_2—$CONH_2$	128.13	
谷氨酸或谷氨酰胺		Glx	Z			
甘氨酸	glycine	Gly	G	—NH—CH_2—CO—	57.05	5.97
组氨酸	histidine	His	H		137.14	
异亮氨酸	isoleucine	Ile	I	—NH—CH—CO— CH—CH_3 CH_2—CH_3	113.16	6.02
亮氨酸	leucine	Leu	L	—NH—CH—CO— CH_2 CH—CH_3 CH_3	113.16	5.98
赖氨酸	lysine	Lys	K	—NH—CH—CO— $(CH_2)_4$ NH_2	128.17	9.74
甲硫氨酸	methionine	Met	M	—NH—CH—CO— $(CH_2)_2$ S—CH_3	131.19	5.74
苯丙氨酸	phenylalanine	Phe	F		147.18	5.48
脯氨酸	proline	Pro	P		97.12	6.30
丝氨酸	serine	Ser	S	—NH—CH—CO— CH_2OH	87.08	5.68
苏氨酸	threonine	Thr	T	—NH—CH—CO— CHOH CH_3	101.11	6.16
色氨酸	tryptophan	Trp	W		186.21	5.89

续表

氨基酸	英文名	三字母符号	单字母符号	结　构	残基分子质量/Da	等电点
酪氨酸	tyrosine	Tyr	Y		163.18	5.66
缬氨酸	valine	Val	V		99.13	5.96

除了表 1.1 列出的 20 种基本氨基酸外，实际上还有几百种不常见的氨基酸参与某些蛋白质的组分，可称为稀有氨基酸；另一些氨基酸出现频率要高些，如羟脯氨酸、甲基组氨酸和甲基赖氨酸、γ-羧基谷氨酸等，但它们是在蛋白质生物合成后转变而来的，因此不计入基本氨基酸之列。

蛋白质的化学结构如下：



这是一个链状的结构，经水解后，将肽键断裂，转变成一系列的氨基酸。链中相当于氨基酸的单元结构称为残基。而上述链称为肽链。肽链的氨基一端称为 N 端或氨基端，羧基一端称为 C 端或羧基端。习惯上将 N 端列在左边，C 端列在右边。

从氨基酸的结构看，半胱氨酸的侧链上有巯基(—SH)。巯基的稳定性较差，在碱性 pH 下易氧化成硫—硫键(或称二硫键)。如果一条肽链上的两个半胱氨酸的巯基形成二硫键，肽链形成链内二硫键。假如两条肽链间形成二硫键，就出现由两条肽链组成的蛋白质(如胰岛素)。蛋白质分子有两对以上的二硫键，或者是蛋白质分子有一对二硫键和一个半胱氨酸，二硫键可能有不同的连接方式，于是出现了异构体。异构体的数目因半胱氨酸含量不同而异，但是在天然的蛋白质分子中没有这种异构体，只有一种连接方式。而在重组蛋白质的复性中，可能出现二硫键的错配对。

肽链由许多氨基酸通过肽键(相邻两个氨基酸脱去一分子水)连接而成。20 种氨基酸前后沿一定的排列次序形成特定的结构。当氨基酸组成及它们的排列次序和长短不同时，就形成了不同的蛋白质；或者，不同的蛋白质有特定的氨基酸序列，这种序列是蛋白质化学结构的最重要的内容。

蛋白质分子中还有非肽链结构的部分，有些也通过共价连接(糖基化、磷酸化等)，有一些是通过非共价结合的(辅基或配基)。

根据蛋白质结构有不同的层次，可分别称之为蛋白质的一级结构、二级结构、三级结构及四级结构。

一级结构是指肽链中的氨基酸排列序列，二硫键的定位也是一级结构的重要内容。下面是猪胰岛素 A 链的一级结构的两种表示形式：

Gly-Ile-Val-Glu-Gln-Cys-Cys-Ala-Ser-Val-Cys-Ser-Leu-Tyr-Gln-Leu-Glu-Asn-Tyr-Cys-Asn

G I V E Q C C A S V C S L Y Q L E N Y C N

蛋白质的二级结构也可称为构象单元，因为它们是蛋白质复杂的空间构象的基础。常见的构象单元有：α-螺旋；β-折叠；β-转角；无规则卷曲。一些构象单元的结构不是不可改变的，pH、温度、溶剂极性等环境因素可以影响单元的变化。

具有二级结构的肽链在空间走向，形成具有空间三级结构的蛋白质，如所谓球蛋白类，它们几乎都有生物功能和活性，其中有些还能进一步相互作用，形成更高层次的结构。蛋白质之所以有功能，是因为分子表面有可以进一步作用的基团，不同的蛋白质由于分子表面结构不同，可作用的基团分布和组合也不同，这种特定的结构就是各种蛋白质具有不同的功能和活性的分子基础。

四级结构可以认为是一些特定三级结构的折叠肽链通过非共价键而形成的大分子体系，也可称为亚基的装配(assembly)。装配是一个非常复杂的过程，它能使各亚基间相互调节，使蛋白质分子的功能更完善。

蛋白质的分离纯化是研究蛋白质的基础和起点。蛋白质的来源是多样的，主要来自动物、植物的各种组织和微生物，取材要采用含量丰富的器官，这有时还和季节有关；近年来，还来自基因工程的方法。

抽提蛋白质的第一步是将组织粉碎，破坏细胞，以便于抽提出蛋白质；抽提液的选择也因蛋白质而异。一般用低浓度的缓冲液，有时在从肌肉抽提时也用稀碱，因肌肉中含有乳酸，最后抽提液的 pH 却是中性的。在提取膜蛋白或包含体内的蛋白质时，还加入(非离子型)表面活性剂或变性剂以增加溶解度。各种细胞中普遍存在各种蛋白水解酶，会导致蛋白质的降解，低温操作和加一些相应蛋白酶抑制剂(如 DFP、PCMB、PMSF 等)、螯合剂(EDTA 等)是有益的。

将蛋白质抽提出来后，就要进一步纯化。纯化不外乎两方面：一是将大体积变成小体积；二是把多组分蛋白质只留下单一种的蛋白质。前者有吸附法、超滤法[1]、冻干等。后者有根据蛋白质的分子形状和分子质量大小、电离性质、溶解度和疏水性、生物功能专一性等，常用有各种盐析法[2]、超离心沉降、结晶[3]、电泳、凝胶过滤，离子交换色谱、亲和色谱等。近年来发展的 FPLC 和 HPLC 大大促进了许多蛋白质在毫克级的纯化，其量足够用于蛋白质的化学研究。一些新型无载体的液相电泳的发展[4,5]，能在数十毫克到数克水平上制备蛋白质，并开拓了一个新领域。精制蛋白质的技术将在本书有关章节中阐述。

第二节　蛋白质的纯度

蛋白质的纯度是一个重要的指标，尤其在重组蛋白的质量控制中。但它也是一个相对的指标，而纯度要求 95%、99%或 99.9%需根据实际要求而制定。正如化学试剂有化学纯、分析纯、光谱纯等，蛋白质的纯度一般指是否含有其他杂蛋白，而不包括盐、缓冲液离子、十二烷基硫酸钠(SDS)等小分子在内。一些蛋白质在硫酸铵糊中相当稳定，一些酶

在甘油中保存则很稳定。而在遗传工程产品中，却要考虑这些小分子的存在与否。

蛋白质纯度如用百分数表示，有时还和检测方法、定量方法、所选用仪器的参数有关。例如在 HPLC 分析蛋白质纯度时，一般要求以 280nm 检测，这是蛋白质的吸收高峰，而很可能不是非蛋白质杂质的紫外吸收峰，甚至小分子物质在 280nm 处没有吸收，因此都以 280nm 的光吸收来定量处理，检测不到非蛋白质杂质和小分子物质的存在，表明所得的蛋白质纯度会偏高。

在色谱中普遍采用积分面积的方法，这也不是很妥当的。而且在用计算机或积分仪在线检出时，仪器的参数设置与所得的报告很有关系，设置不妥往往导致结果不够准确。

目前常用的鉴定蛋白质纯度方法有：①聚丙烯酰胺凝胶电泳和十二烷基硫酸钠-聚丙烯酰胺凝胶电泳（SDS-PAGE）；②毛细管电泳（CE）；③等电聚焦（IEF）；④高效液相色谱（HPLC），包括凝胶过滤、各种反相色谱、离子色谱、疏水色谱等；⑤质谱（MS）。

一些新的有效方法也在引入分析蛋白质的纯度，如质谱等。由于它测定分子质量的精确度可达万分之一，因此丢失一个氨基酸残基（平均分子质量约 110Da）就可以检出；如有氨基酸的取代形成突变株（可能分子质量有几十道尔顿的差异），也非常容易被发现。

此外，还有一些化学法，例如观察末端或 C 端是否均一，也是相当有效的方法，有时甚至比常用的 HPLC 更有效。比如有一个产品在 HPLC 中已是一对称的峰，该生产单位认为是 100%纯，但在蛋白质的 N 端分析中指出纯度还不到 90%，有一定量的其他蛋白质存在。

用末端分析法检查蛋白质的纯度，在国内目前还没有引起重视，其实在测 N 端序列的同时，应该也给出 N 端均一性（蛋白质的纯度百分数）的数据。

在遗传工程产品的鉴定中，又常常发现 N 端是不均一的，存在着微观不均一性。例如新型白细胞介素－2（IL-2，Ser-125）的 N 端，部分是 Ala-Pro-Thr-Ser-Ser……，部分携有 Met，故其序列是 Met-Ala-Pro-Thr-Ser……但一般不影响其生物活性，仍认为是均一的。但美国 Hoffmann-La Roche 同类产品均为 Met-Ala-Pro-Thr-Ser……

按照严格的要求，用电泳法证明蛋白质的纯度，在一种 pH 下电泳说明该蛋白质是纯的，这是不够充分的。应该取两种 pH 缓冲液，它们分布在蛋白质等电点的两侧，在这两种 pH 下电泳都证明此蛋白质是纯的，这样才可靠。

应该指出，只用一种方法鉴定蛋白质纯度是不够的，至少应该用两种以上的方法，而且是两种不同的分离机理的方法来判断蛋白质纯度才比较可靠。因为常常发现某一样品在凝胶过滤中是纯的，而在离子色谱中却分辨为两个组分，反之亦然。

又如，一种样品用凝胶过滤法和 SDS－PAGE 电泳证明是纯的，但这仍不充分，因为这两种方法的机制是相同的。

第三节　蛋白质的定量

测定蛋白质的量是研究蛋白质的最基本的一步，而对蛋白质定量方法的原理、适用范围、灵敏度和可靠性程度有一些了解是必要的。

往往一些单位在送蛋白质/多肽样品进行氨基酸组成分析或进行其序列分析时，样品

实际浓度和所报(预期)的浓度有很大的差异,甚至有数量级的差异,以致上样品量太少而无法分析,造成人力、时间、费用上的浪费。

更有甚者,根据一些片面数据,没有测定蛋白质浓度,进行光谱和定量氨基酸分析,就断言所得的样品是蛋白质。因此,蛋白质的定量看来是“小事”一桩,但却是最基本的要事,如不重视,会在根本上出问题。

溶液中蛋白质的浓度测定方法很多,最早的经典方法是克氏定氮法,因为每一种蛋白质都有其恒定的含氮量(14%～16%)。其方法是将一定量的蛋白质用浓硫酸消化分解,使其中的氮变成铵盐,再与浓 NaOH 作用,使氨气放出而被吸收于标准酸液中,用反滴定法滴定残余的酸,或用硼酸吸收后,再用标准酸直接滴定,求得该蛋白质样品中的含氮量[6,7]。

此法虽有普遍性,但操作繁琐,目前应用较少。稍后应用较广泛的方法是双缩脲法、福林-酚法和紫外吸收法,近年来大多数人用考马斯亮蓝法。

选择哪一种方法,一般考虑的原则是准确、操作方便、影响因素少。

一、光吸收法

蛋白质在 280nm 左右有吸收高峰,这主要是由于蛋白质中的芳香族氨基酸(Trp 和 Tyr)的缘故,因此可以用 280nm 光吸收值来定量溶液中的蛋白质。但是每种蛋白质中的芳香族氨基酸含量是不同的,而且有较大的差别。最简单的方法是采用 280nm 光吸收为 1 时等于 1mg/ml 来计算。这样简单的处理在准确性上有不足之处,但其测定时间短,用量极少,而且不消耗样品,故被广泛采用。

考虑到蛋白质样品中可能含有少量核酸类杂质,因此采用 280nm 和 260nm 光吸收以下式计算。

$$\text{蛋白质浓度(mg/ml)} = 1.45A_{280}\text{nm} - 0.74A_{260}\text{nm}$$

最准确的方法是用蛋白质的摩尔消光系数计算。在蛋白质纯化后,将此纯的蛋白质对水充分透析,然后冻干或在真空干燥器中干燥,继而在 105℃ 下恒重,精确称量 1～10mg,再定量溶于一定体积中(通常为 1mg/ml),测量 280nm 下的光吸收,从而得出该蛋白质的摩尔消光系数。这种方法最为方便、准确,适于微量测定。

二、双缩脲法

此法是基于蛋白质之肽键—CO—NH—有双缩脲反应,在碱性溶液中与 Cu^{2+} 络合显蓝色[8]。该法受蛋白质之特异性影响较小,适用于较大量蛋白质(毫克级)的测定。

三、福林-酚法

此法灵敏度和准确性都较高,但操作较繁琐,此外,影响因素(譬如去垢剂干扰这一方法)也较多[9]。

Smith[10]提出一种新的 BCA 测定蛋白质的方法,灵敏度和重复性均佳,受干扰少,Pierce 公司有商品供应。

四、考马斯亮蓝 G-250 法

此法适合于定量 10～100μg,操作简单,应用广泛。考马斯亮蓝 G-250 又叫"home-made"试剂[11]每个实验室可以方便地自行配制。

五、氨基酸分析法

此法很准确,最接近真实值,因为其他方法都用一个已知的纯蛋白(大多数情况下采用牛血清白蛋白,BSA)作一标准曲线,将待测蛋白与之相比,实际上两种蛋白质的芳香族氨基酸含量不完全一样,所以会产生系统误差。而用氨基酸分析来计算是测得该蛋白真实值,但稍偏低一点,因为蛋白质在酸水解时,Trp 被破坏,Cys 也遭受很大破坏,Ser 和 Thr 回收也在 90%左右。

上述方法的操作细节可参见《蛋白质化学研究技术》[7]和《蛋白质和酶学研究方法》[12]。

第四节　蛋白质的脱盐(除去盐和非共价结合的小分子)

蛋白质溶液中的大多数盐和低分子质量的物质可以对 0.1mol/L NaCl 透析而除去,然后再对水充分透析。如果用过甲酸处理过的样品,由于缺少卤原子,所以先用其他盐或缓冲液透析,如挥发性的乙酸铵、稀的乙酸或碳酸氢铵,然后再对水透析或直接冻干。

用凝胶过滤法(Sephadex G-25 或类似的介质)和 HPLC 的凝胶柱脱盐。应用这类技术时首先要考虑蛋白质的回收,特别对少量蛋白质,尤其是碱性蛋白质,则常常因载体对蛋白质的吸附作用而导致回收差。此时的流动相一般不能用水,而推荐用 0.1mol/L 醋酸或 0.1mol/L 碳酸氢铵,因为可以在冻干过程中除去这些挥发性的盐。

在蛋白质纯化的最后阶段,常用挥发性溶剂(例 0.1% TFA/ACN)的反相 HPLC,则可得到无盐的蛋白质样品,在我们实验室常用此法于蛋白质序列分析中净化样品。如果对某些蛋白质的回收低,建议用短柱(例 PE-ABD 公司的 RP-300 柱,30mm×2.1mm)较为理想。

比 HPLC 更简单的方法是用 C-18 的 Sek-Pak,或用 Millipore 公司的超滤离心过滤管,但耗费较大。

如果蛋白质是用 SDS-PAGE 纯化的,则非共价结合的分子往往可以有效地除去,但又引入了 SDS 和其他一些分子(Tris,Glycine 等)。从 SDS-PAGE 电泳的凝胶中回收蛋白质的方法很多,但不尽如人意。如采用电印迹,往往很有效。因为 PVDF 膜结合蛋白质的亲和力大大高于盐、去垢剂等小分子物质,从而达到蛋白质脱盐的目的[13,14]。

Stone 推荐用 TCA 沉淀蛋白质而脱盐,蛋白质浓度>100μg/ml(如果蛋白质浓度太低,得不到沉淀,这时可用真空离心浓缩蛋白质的方法),加总体积 1/9 的 100% TCA,

TCA 最后浓度为 10%,在冰浴放置 30min 后离心收集沉淀物,用 100μl 冷丙酮洗两次,气流干燥[15]。

Stone 还建议用丙酮沉淀法,1ml 以下的蛋白质对 0.05%SDS,5mmol/L NH_4HCO_3 充分透析,然后转移到 1.5ml Eppendorf 管中真空离心干燥;加 50μl 水;再加 450μl 1mmol/L HCl 的酸性丙酮,在 −20℃ 放置 3h,离心收集沉淀,用 100μl 冷丙酮洗沉淀两次,气流干燥[15]。

注意:如果 SDS 的量大于 0.05%,则 SDS 易与蛋白质形成共沉淀。

金属离子或辅酶往往和蛋白质紧密结合,这可将蛋白质变性后用透析法除去。

第五节　蛋白质的分子质量测定

蛋白质分子质量的测定早期采用超离心分析和光散射法,这需要较高级的专用仪器和较多量的蛋白质的样品,目前应用很少。

二十多年前采用凝胶过滤法测定蛋白质的分子质量[16],应用最广泛的是 Sephadex 列(G-75、G-100)等。近年来由于解决了分离介质的刚性问题,有了 HPLC 的凝胶过滤系统(例如 Waters 的 I-60、I-125、I-250 和 Beckman 或 Toyo Soda 的 TSK2000SW,3000SW,4000SW),这样测定分子质量只需 1h 即可。

但在一般实验室应用的常规方法是 SDS-PAGE[17],原理是在 SDS 存在条件下,蛋白质表面都携带了大量负电荷,呈杆状分子,在此情况下,不是根据蛋白质的电荷而是根据其分子形状和大小来进行分离。用量为 0.1~1μg,这种方法误差为 5%~10%,但极为简便。

凝胶过滤是测定完整蛋白质分子的分子质量,而 SDS-PAGE 是测定蛋白质亚基的分子质量,同时用这两种方法测定同一蛋白质分子质量,可以方便地判断样品蛋白质是否寡聚蛋白质。

较新方法是用毛细管电泳(凝胶筛分或无胶筛分)测定蛋白质的分子质量,仅需纳克量,而且还可用积分仪或电脑联机精确定量。同时此法对蛋白质的分辨率和准确性优于 SDS-PAGE 方法。在某一批 IL-3 产品的分子质量测定中,用 SDS-PAGE 得到均一的区带,但在毛细管筛分电泳中为两个毗邻的峰。两个峰的分子质量也有约 1 000Da 的差异,我们进一步用质谱验证了这一结果。

最新的方法是用质谱测定蛋白质的分子质量。20 世纪 80 年代初期,质谱用于测定分子质量的有快原子轰击质谱(fast-atom bombardment mass spectrometry, FABMS)和等离子体解吸飞行时间质谱(plasma desorption time-of-flight mass spectrometry)。

近年来,高分辨率的磁质谱可精确测定分子质量 2 000Da 以下的多肽。到 90 年代,由于电喷雾质谱(ESI)有了长足的发展,可以用于测定分子质量约 5 万 Da 的蛋白质,而且只需皮摩尔(pmol)量的蛋白质,且其精度可达 0.01%。而 MALDI-TOF 则可测定蛋白质分子质量高达二三十万道尔顿。图 1.1 所示为 ESI 质谱测定细胞色素 c 的分子质量,右角是其电荷分布图,据此计算出分子质量。细节可参见相关章节。

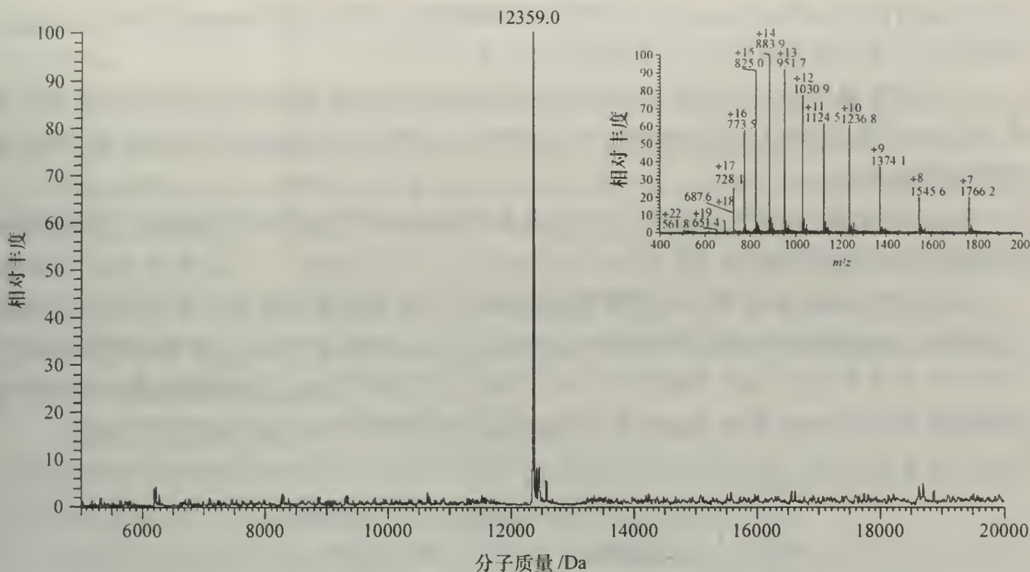


图 1.1　电喷雾质谱法测定细胞色素 c 分子质量

第六节　蛋白质的等电点测定

等电聚焦法可以测定蛋白质的等电点,同时又能鉴定蛋白质的纯度,是一种一举两得的方法。严格讲,一种蛋白质只有一个等电点,不同研究小组报道同一种蛋白质的等电点有所差别,可能和所用的具体方法不同有关。如我们实验室测定重组人表皮生长因子

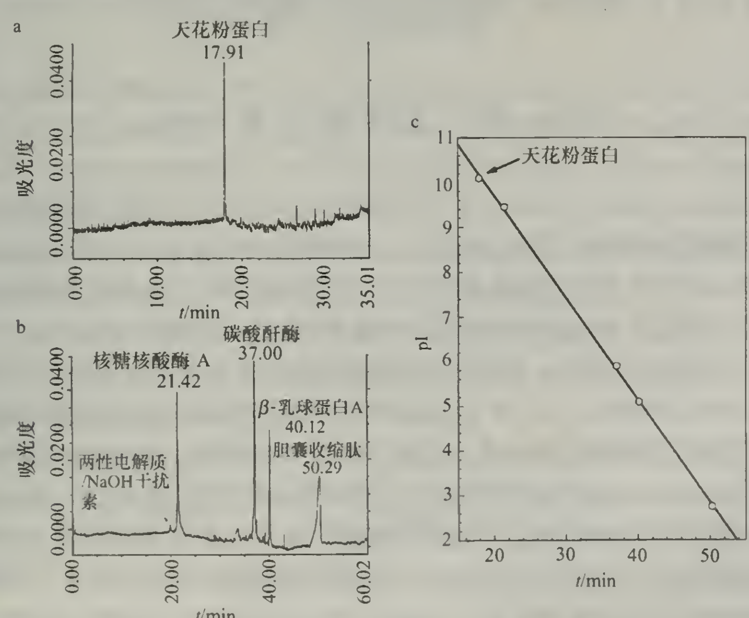


图 1.2　CIEF 测定天花粉蛋白的等电点

a. 天花粉蛋白的等电点测定;b. 标准蛋白质的等电点测定;c. 左侧二图实验值的作图

(rhEGF)时,用 Bio-Rad Mini-IEF 在聚丙烯酰胺凝胶上测得它的等电点为 4.6[18],而用全液相的毛细管等电聚焦(CIEF)测得其等电点是 5.0[19]。

人们常常遇到某一产品用一些方法证明是纯的,但在等电聚焦分析时出现较多的条带,对此众说纷纭,尚无公认的解释。有的专家认为可能是空间构象不同而引起,但还有待实验来验证。

CIEF 测定蛋白质等电点的另一个优点是能测定偏碱性的蛋白质等电点,比如天花粉蛋白的等电点被测定为 10.10[5]。

最近,我们实验室发展了毛细管等电聚焦电泳-电喷雾质谱技术(CIEF-MS),这种技术的等电点分辨率极高(能分辨的等电点差异小于 0.04pH 单位)。一些标准样品用这种方法分析可发现亚型(分子质量完全一样,等电点不同)的存在,其中细胞色素 c 两种亚型的等电点分别为 9.60 和 9.55,差异只有 0.05,一般的方法根本无法分辨(图 1.3)。

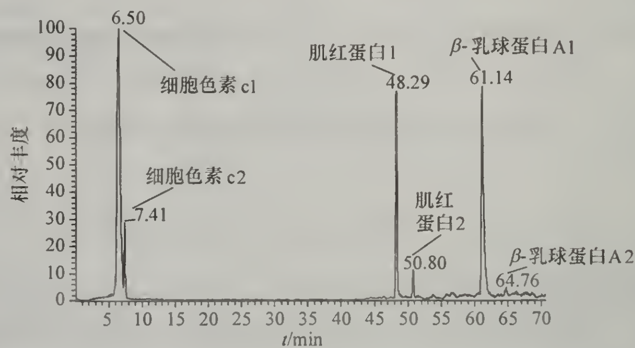


图 1.3　CIEF/ESI-MS 联用分析标准蛋白混合物
(使用 Pharmalyte 3.5－10)

第七节　氨基酸定量分析

自从 Spackman、Stein 和 Moore 建立氨基酸自动分析方法以来,无论在方法学、仪器自动化或高灵敏度方面都有了很大的发展。氨基酸分析可以分为柱后反应法和柱前衍生法两大类。前一种方法是将游离氨基酸经过色谱柱分离后,氨基酸再与显色剂(茚三酮、荧光胺、邻苯二甲醛)作用,仪器需用专门的氨基酸分析仪或用 HPLC 加上柱后衍生装置。这种方法比较稳定,因为一些可能有影响的“杂质”在柱分离过程中已与氨基酸分开,然后再各自与显色剂作用。后一种方法是将各种氨基酸和荧光试剂先作用生成氨基酸的衍生物,然后再用柱把各衍生物分离,直接检测衍生物的荧光,这种方法的特点是灵敏度高,检出极限已达 1 pmol,且可以利用 HPLC 进行氨基酸分析,譬如 OPA 技术。缺点是荧光强度依赖时间变化,稳定性欠佳,要求控制很严格[20]。近年来报道的 PITC 是一种很有效的柱前衍生技术,并且可以在 254nm 下检测,操作简便,作者也作了一些尝试[21]。

氨基酸的水解方法通常采用在 5.7mol/L HCl 真空状况下 110℃ 水解 24h,也有在 150℃ 下快速水解 4h。不同条件下,各种氨基酸的回收率会有所不同,而且色氨酸会受到破坏,通常采用 3mol/L 巯基乙磺酸或 4mol/L 甲磺酸来防止色氨酸被破坏。一种新的方

法是在水解样品中加入保护剂十二烷硫醇(dodecanethiol)和EDTA,据称色氨酸的回收率可达80%。

Stein实验室引入气相水解法,操作方便,样品损失少,回收率也高。目前已有把蛋白质水解、自动进样、氨基酸分析和定量报告联成一个系统的氨基酸分析仪商品出售。

在酸水解条件下半胱氨酸的定量很不准确,偏低严重。通常在酸水解蛋白质前用过甲酸氧化成磺基丙氨酸(cysteie acid);或先还原蛋白质,用碘代乙酸处理把半胱氨酸转变成羧甲基半胱氨酸。一种较新的方法是把蛋白质用DTT还原后,加入乙烯吡啶(4-vinylpyridine),然后分析其衍生物。也有报道用DTDPA(3,3′-Dithiodipropionic acid)与半胱氨酸和胱氨酸形成Cys-MPA,此衍生物能在HPLC中和其他氨基酸衍生物分离,故可用于准确定量半胱氨酸。

如果采用电泳法(例如SDS-PAGE)分离到样品,则很难从凝胶中洗脱下来,可采用点印迹法把样品转移到PVDF膜上,然后在PVDF膜上直接进行盐酸水解和氨基酸分析称之为原位分析(*in situ*),在膜上分析氨基酸的误差要高于常规方法。

人们在氨基酸分析时重视精确性和重复性,而常常忽略了其准确性。为了较好地评价氨基酸分析的准确性,过去作者用胰岛素为标准样品来判断分析方法的可靠性。近年来采用"测试肽"的方法,即将20种常见的氨基酸人工合成为一个20肽。然后用酸水解,每种氨基酸残基出现频率都是1,易于校正回收。

氨基酸分析在基因工程产品分析中是一项重要的辅助指标。往往有人将产品做了氨基酸分析,见到有氨基酸的存在,就断言此产品为蛋白质,这是很不充分的。在此要强调氨基酸的定量分析,究竟取多少样品量去水解和做氨基酸分析?如果取10μg样品去分析,最后各种氨基酸总和才1μg,那么尽管该样品有生物活性或疗效(如果确有的话),要考虑可能该活性不一定是来自蛋白质,因为在该产品中,蛋白质只占10%,算是一种"杂质"。

进行重组蛋白质的氨基酸分析,可以和标样比较,以确认重组蛋白质的氨基酸组成是否和天然蛋白质的氨基酸组成一样。如果是某种未知蛋白质,可以到蛋白质数据库(Protein Data Bank)去查阅,看看和哪一种已知蛋白质的组分相似,再作进一步确认,或者此样品就是一种新的蛋白质。

对某单位生产的γ干扰素进行分析,用BrCN裂解后找不到C端肽,同时观察到氨基酸分析中精氨酸明显缺少(γ干扰素的133个氨基酸中精氨酸占8个,其中5个集中在近C端),提示C端可能是不完整的。由于氨基酸分析提供重要信息,从而让工作深入下去,最后证明C端有所缺失,少了约13个氨基酸的肽,但没有危及生物活性[22]。当时如有质谱技术,解决此类问题会简单得多。

在基因工程的重组蛋白质中,分子质量一般在10 000～20 000Da之间,氨基酸分析的准确性有多大?实际意义有多少?这个问题上是有不同看法的。有些产品上报的氨基酸组成和理论值完全相同,反而是不太合理的,因为一些羟基氨基酸(丝氨酸,苏氨酸酸水解时有所破坏;又因为各种肽键在酸水解时,有些肽键水解不完全,有些肽键受到破坏,因此和理论值总有些差异却是正常的现象。一般讲,目前很少有人在科学研究论文中将蛋白质的氨基酸组成作为一项数据来发表,因为往往和最后的蛋白质序列有较大的差异,还不如等待序列的数据为好。

氨基酸分析在肽谱和点突变肽的分析中是很重要的，同时也可以是很精确的。因为这种肽通常是含 5～20 个氨基酸残基的肽，又常常是以特定的某一个酸性或碱性氨基酸为 C 端。从氨基酸定量分析而言，分析 20～30 个氨基酸残基肽的精确性要比含 100～200 个氨基酸残基的蛋白质的精确性大得多。其次，酶解肽段（最常用胰蛋白酶）可用一个碱性氨基酸作基准 1 来计算，这样准确性又可有所提高。在新型 IL-2 的制备工作中（Cys-125 被 Ser 或 Ala 所取代），利用二阶微分光谱法[23]可以方便地判别点突变的肽，然后用氨基酸分析或质谱，或序列测定来进一步确认。

第八节 肽 谱

肽谱技术是指蛋白质被酶解或化学降解后肽段的分离分析或制备。

肽谱分析最早称之为 fingerprinting（指纹术），用在血红蛋白 A（Hb A）和血红蛋白 S（Hb S）的分子病研究中，即将 Hb A 和 Hb S 分别用胰蛋白酶酶解，然后将它们的肽分别进行纸电泳，再转动 90°进行纸层析。Ingram 用此法观察到 Hb A 和 Hb S 之间有一个肽斑点上的差别，进一步的肽顺序分析指出仅仅是一个氨基酸残基（一个净电荷）的差别，正常的 Hb A 中是 Glu（β-6），异常的 Hb S 中为 Val（β-6）。该技术发展到现阶段，肽谱（peptide-mapping）用 HPLC[22,23]和 CE[24]来做，HPLC 主要是用 RP-HPLC 根据肽的长短和疏水性质来分离，如果肽的亲水性强，则柱不能滞留住这些亲水性强的肽段，达不到分辨开的效果，如果某些肽的疏水性很强，往往会粘在柱上洗脱不下来，而 CE 有时可以避免这些弊端。

早期对遗传工程产品的肽谱分析指定用 SDS-PAGE 这一方法。用 SDS-PAGE 来进行“肽谱”的分析，有方便的一面，但也有其局限性，譬如用溴化氰化学裂解蛋白质时，生成的肽段较大（文献上常称为“fragment”，而不称为“peptide”），由于肽段少，SDS-PAGE 可以根据它们的分子质量不同而分辨开，但对于小分子的肽，往往无法分辨或在染色脱色过程中丢失。我们用溴化氰裂解肌红蛋白时，它的小肽（分子质量 2 512Da）在 SDS-PAGE 和电印迹后回收仅 10%左右（基于氨基酸序列分析），再小的肽则更难操作，故此法有其局限性。同理，更不适用于酶解后肽谱的分析，因为酶解的小肽段往往更多。

在用 BrCN 化学裂解 IL-2 后，再用 SDS-PAGE 进行肽谱分析，由于 IL-2 中有 4 个 Met，理论上应有 5 个肽段，实际上 SDS-PAGE 肽谱上仅见到 3 个肽段的区带，因为有 2 个肽段分子质量小，在染色和脱色过程中丢失。在这种情况下，最好用 HPLC 分析肽谱。

如果 Met 在蛋白之中分布较匀，得到的肽段大小差不多，都能在 SDS-PAGE 中检出，看来“功德圆满”了，但 SDS-PAGE 测分子质量不是很精确，最好是将此肽谱印迹（blotting）到 PVDF 或 ProBlot 膜上，再置于蛋白质序列仪上做几个 N 端序列，以确认是哪一段肽。

我们在 IL-3 的肽谱工作中，采用 TPCK-胰蛋白酶酶解，HPLC 进行肽谱分析和收集各个肽峰，对每一峰作了氨基酸组成分析和质谱分析，IL-3 在酶解后理当得到的肽都在肽谱中得到验证，这样的结果是比较有说服力的（图 1.4，表 1.2）。

在没有条件验证的情况下，报批的三批样品肽谱相同，这不够充分，最好有标准样品作对照，三批报批样品的肽谱和标样的肽谱一样，这样安全系数就大多了，以防万一由于某种原因，三次有误而且能重复，得出不恰当的结论。

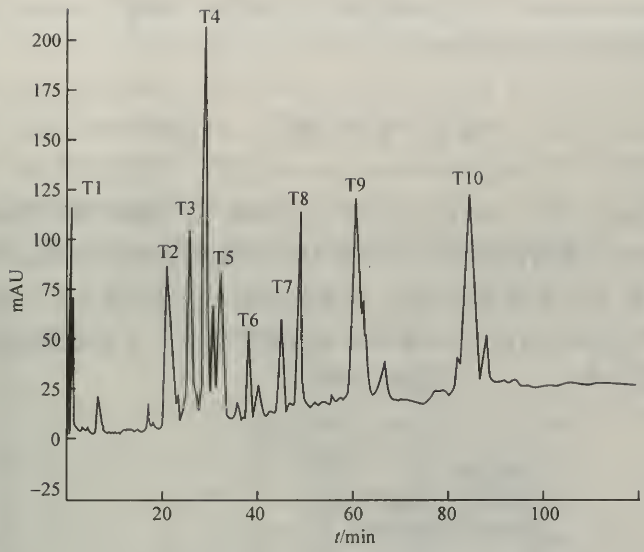


图 1.4　IL-3 的胰蛋白酶酶解肽谱

表 1.2　IL-3 的胰蛋白酶酶解肽段的氨基酸分析

	T1	T2	T3	T4	T5	T6	T7	T8	T9	T 10*
Asp				2.5	1.6		1.1	1.3	7.0	4.1
Thr			2.7			0.9		2.8		3.6
Ser							2.9	1.5		2.0
Glu			1.3	1.0	1.0		2.2	3.7	4.4	1.4
Pro		1.0	2.0		0.8				2.3	1.8
Gly				1.0					1.4	
Ala	0.9		1.2		1.0		2.0	3.0		2.9
Cys										
Val	0.7									0.8
Met			1.0						1.0	1.0
Ile		1.8					2.0	1.1	0.8	2.1
Leu			1.3		1.0	1.9	2.1	3.2	6.3	5.0
Tyr						0.9				
Phe					1.1	1.0		0.9	1.2	
His		1.8		1.0						1.2
Lys	1.4	1.2	1.3			1.1	1.1			1.0
Arg	0.95			1.0	0.8				1.3	1.2
Trp										
No.	64～66	95～100	1～10 N端肽	101～108	56～63	111～116	67～79	117～133 C端肽	29～54	11～28& 80～94

* 肽 10 是两段肽，因为酶解前没有用 DTT 还原二硫键，所以这两段肽由二硫键连接在一起，反之也证明 IL-3 中二硫键的配对是正确的。

近年来，随着 LC-MS 和 CE-MS 的普遍应用，使收集到的肽峰立刻得到精确分子质量的测定，这样使肽谱的可靠性大大增加。

此外，“肽质量指纹”(peptide-mass fingerprinting，PMF)的出现，即用 MOLDI－TOF

质谱直接分析蛋白质的酶解产物，多肽出现先后是按分子质量大小排列，这对于已知序列的重组蛋白质样品的分析更加有利。

第九节 质 谱

在分子质量测定一节中，我们已初步涉及到质谱(MS)测定分子质量的精确程度；在图 1.5 中，我们见到人胰岛素(B-30 Thr，Mr 119)和猪胰岛素(B-30 Ala，Mr 89) 的差别仅在一个氨基酸残基，分子质量净差 30Da，质谱能准确无误地分辨它们。因此，质谱用在重组蛋白质的质控中是相当有效的，如果在表达过程中丢失一个氨基酸残基，至少在分子质量上会少 75Da(最小是 Gly)，肯定能鉴别出来。

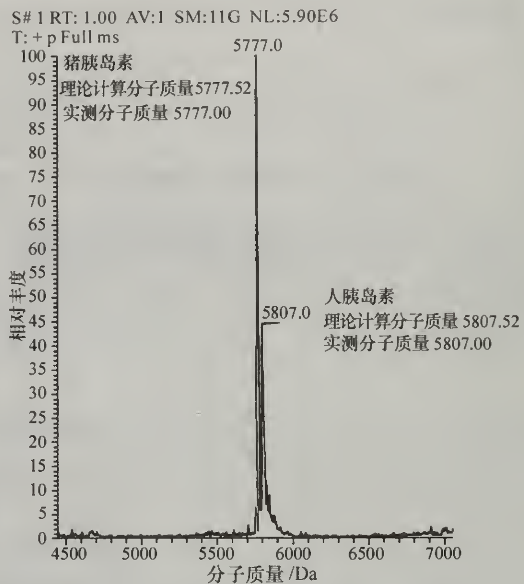


图 1.5 人胰岛素和猪胰岛素的质谱图

质谱在 90 年代的发展，特别是两种软电离技术的发展，使生物质谱在质谱界成为主要角色。蛋白质组学又为生物质谱提供了一个发挥才华的大舞台。它们中首选的是 MALDI-TOF，它的大容量分析、单电荷为主的测定分子质量高达 30 万 Da、干扰因素少，适合蛋白质组的大规模分析之需。其次 ESI 为主的 LC-MS 联机，适于精细的研究。

目前，质谱广泛用于蛋白质的纯度鉴定、分子质量测定、序列测定、肽(质)谱分析、二硫键、乙酰化、糖基化、磷酰化，以及非共价结合等研究中，成为研究蛋白质不可缺少的技术，具有广阔的发展前景。以往的美国质谱年会，参加者约千余人，近年来为三四千人，其中很多是生化学家，这和 ESI 和 TOF 在生物学中的广泛应用是分不开的。质谱年会上生物学和医学方面的论文也由 10%上升到近 40%。

对于蛋白质，质谱发展还有一个重要的功能是肽的测序，这是一个被认为很有前途的测序方法，称为串联质谱(Tandem-MS)，即在第一级质谱得到肽的分子离子，选取目标肽的离子作为母离子，与惰性气体碰撞，使肽链中的肽键断裂，形成一系列子离子，即 N 端碎片离子系列(b 系列)和 C 端碎片离子系列(y 系列)，将这些碎片离子综合分析，可得出

肽段的氨基酸序列。在以四级杆为质量分析器的质谱，通常是将三个四级杆串联起来，实现二级质谱，即 MS/MS，这就是所谓三级四级杆质谱。在以离子阱为质量分析器的质谱中，采用离子阱可多次捕捉母离子和子离子，因此可实现多级质谱，即 MSn，一般可达五级以上，对于分子结构分析很有用处。目前的测序方法基本上是在 ESI 质谱上发展起来的，MALDI-TOF-MS 最近发展了 PSD(post source decay)技术，在一定程度上可进行序列分析，但不如 ESI-MS 测序应用广泛。

图 1.6 是我们实验室对一类新药 TPO 用质谱进行的全序测定。

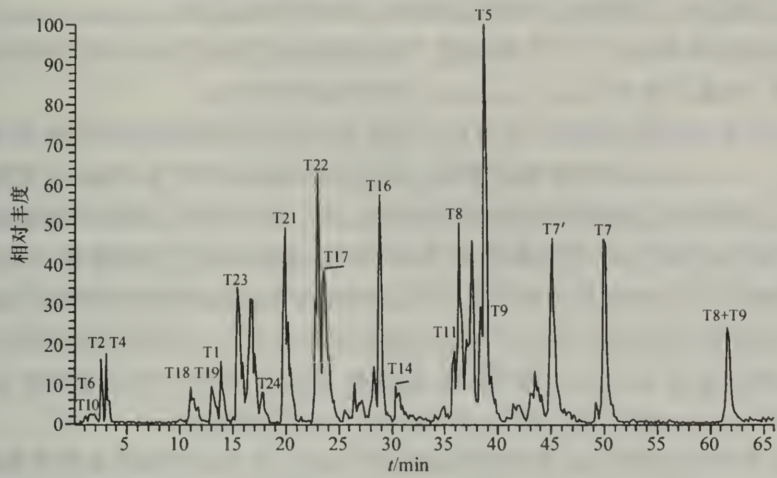


图 1.6　TPO 的质谱测序

TPO 的氨基酸序列

```
1                    10                   20                   30                   40
S-P-A-P-P-A-C-D-L-R-V-S-L-K-L-L-R-D-S-H-V-L-H-S-R-L-S-Q-C-P-E-V-H-P-L-P-T-P-V-L
|----------T1----------| |--T2--| |-T3-| |---------T4--------| |----------------------------------T5--
41                   50                   60                   70                   80
L-P-A-V-D-F-S-L-G-E-W-K-T-Q-M-E-E-T-K-A-Q-D-I-L-G-A-V-T-L-L-L-E-G-V-M-A-A-R-G-Q-
-------------------------| |------T6------| |------------------T7------------------------------| |-------
81                   90                   100                  110                  120
L-G-P-T-C-L-S-S-L-L-G-Q-L-S-G-Q-V-R-L-L-L-G-A-L-Q-S-L-L-G-T-Q-L-P-P-Q-G-R-T-T-A-
----------T8-------------------------------| |--------------------T9 --------------------| |--T10--
121                  130                  140                  150                  160
H-K-D-P-N-A-I-F-L-S-F-Q-H-L-L-R-G-K-V-R-F-L-M-L-V-G-G-S-T-L-C-V-R-R-A-P-P-T-T-A-
---| |------------T11-----------------| T12  T13 |--------------T14--------------| |----------T15---
161                  170                  180                  190                  200
V-P-S-R-T-S-L-V-L-T-L-N-E-L-P-N-R-T-S-G-L-L-E-T-N-F-T-A-S-A-R-T-T-G-S-G-L-L-K-W-
---------| |---------- T16-----------------| |------------T17-------------------| |-------- T18------| |---
201                  210                  220                  230                  240
Q-Q-G-F-R-A-K-I-P-G-L-L-N-Q-T-S-R-S-L-D-Q-I-P-G-Y-L-N-R-I-H-E-L-L-N-G-T-R-G-L-F-
--T19------| T20 |-------- T21 ---------| |----------T22---------------| |--------T23---------| |---------
241                  250                  260                  270                  280
P-G-P-S-R-R-T-L-G-A-P-D-I-S-S-G-T-S-D-T-G-S-L-P-P-N-L-Q-P-G-Y-S-P-S-P-T-H-P-P-T-
--T24-----| |----------------------------------------------------------------T25-------------------------
281                  290                  300                  310                  320
G-Q-Y-T-L-F-P-L-P-P-T-L-P-T-P-V-V-Q-L-H-P-L-L-P-D-P-S-A-P-T-P-T-P-T-S-P-L-L-N-T-
-------------------------------------------------------------T25-----------------------------------------
321                  330
S-Y-T-H-S-Q-N-L-S-Q-E-G
-----------------------------------|
```

还有采用氨肽酶或羧肽酶与质谱结合而形成的梯带测序法(ladder sequencing),但由于酶解是一个动态演绎过程,不同氨基酸残基的酶解速度不一,故这个方法有一定局限性。

因此,ESI-MS 可较好地与现有的蛋白质分析方法(如 HPLC、CE)蛋白质测序技术相匹配。在蛋白质翻译后修饰研究中,由于 ESI-MS 强大的 MS/MS 乃至 MSn 功能,可以解析蛋白质的各种修饰的位点,如糖基化、磷酸化等,以及复杂的糖链结构。

值得注意的是,质谱技术目前的发展非常迅猛,无论是 ESI,还是 MALDI-TOF,都在不断改进,以期完善。ESI 着力于提高灵敏度(如采用 nanospray,capillary LC),以及降低盐和去垢剂的影响;MALDI-TOF 则着重于提高精确度,如在离子进入分析器前稍微停留,去除杂质,即延迟抽提(delay extraction)和加强测序功能。

最近 Q-TOF 质谱的发展——MALDI-TOF 由于其仪器的固有性能,决定其在高分子质量区域($m/z>3\ 000$)的质量精度下降,一般认为在高分子质量区域,质量偏差应不超过 100 ppm,否则在肽谱检索时得不到准确的结果。这种现象已被我们实验室的经验所证实。Q-TOF 型仪器,由于其四极杆和飞行时间质谱是以相互垂直方向几何型组成,当离子通过四极杆后,同样的质量而具有不同动能的离子在垂直方向上受到飞行时间质谱的推出电压作用时,通过离子的水平方向动能差异不会对飞行时间质谱分辨率产生影响,因而其分辨率比普通 MALDI-TOF 要高,灵敏度也提高,显然它的质量精度亦有显著提高,其偏差<5ppm。这样在蛋白质肽谱的检索中可提高准确度。

目前,Q-TOF 质谱不但有 ESI 源,也配备了 MALDI 源,这无疑是如虎添翼,具有极高的分辨率和灵敏度。

第十节 蛋白质及多肽的序列测定和末端分析

目前测定蛋白这一级结构的主要方法有:蛋白质化学方法(主要是 Edman 降解法测 N 端和用羧肽酶或化学法从 C 端测起)和从 cDNA 来演绎的方法。其他方法有质谱法和二维核磁共振法。

利用 cDNA 法测定的最长蛋白质是脱脂脂蛋白 B-100(4 532 个氨基酸残基),用经典的蛋白质化学分析法测得的最大的蛋白质是人凝血体系中的 Von Willebrand 因子(2 050 个氨基酸残基)。从蛋白质数据库来看,约一半来自 cDNA,一半来自蛋白质化学方法。虽然 cDNA 技术发展很快,也方便,但目前还不能取代经典的蛋白质化学方法,例如部分顺序测定对识别蛋白质的功能结构域或蛋白质活性部位以及化学修饰部位,还是离不开经典的蛋白质化学方法。

蛋白质的末端氨基酸与在肽键中间的氨基酸不同,蛋白质的一端是以 α-自由基存在的,称之为 N 端,习惯上列在左侧;在另一端是以自由羧基存在,则称之为 C 端,习惯在右侧。1954 年 Sanger 首先把二硝基氟苯应用于胰岛素的 N 端测定,为研究蛋白质的精细化学结构开辟了一个新的领域[25],继后 Edman、Wetman-Libet、Chang 等人相继作了很多贡献。

特别应该强调指出,末端的鉴别能力较之一般的物理化学方法(色谱、电泳等)毫不逊色,有时甚至更为灵敏,所以也是鉴别蛋白质样品纯度的一种好方法。

一、N 端分析

在 N 端测定上，目前最经典的是异硫氰酸苯酯法[26]。自 1950 年代建立后，在 1970 年代发展成一种全自动的测定方法。目前微量液相脉冲蛋白质/多肽自动顺序分析仪，灵敏度约 1pmol，检出极限为 20fmol，加上进行计算机数据处理和分析，既灵敏又方便。但机器有时会出现读/写错误，还是需要有经验的工作者加以最后判断才是最可靠的。

Wittmann-Liebold 发展的双标记法(DABITC/PITC)，用薄层分析鉴定的微量序列测定法不需昂贵仪器，可成为一般实验室普遍采用的技术[27,28]，但需要很熟练的技术才能掌握好。

N 端 15 个序列的测定，很大程度上排除了蛋白质混淆的可能，即不太可能有两种不同的蛋白质在 N 端 15 个序列是一样的。这已是相当常规和公认的。

如果测定蛋白质的几个 C 端序列，从蛋白质的两端来验证，可靠性又增加了许多。再加上用 MS 精确测定蛋白质的分子质量，基本上是不会出差错的。如果有肽谱或肽(质)谱，则是相当完整的鉴定工作。

二、C 端的测定

C 端测定最早是肼解法。我国蛋白质化学前辈纽经义先生在 1955 年就应用此技术测定了烟草花叶病毒蛋白的 C 端[29]。

经典的是用羧肽酶(Cp)的方法，近年来虽用 Cp Y、Cp P，但基本方法仍是一样的，基于不同时间的 C 端氨基酸的释放，是一个动态过程的演绎。先是用 CpA、CpB，近来用 CpY、CpP，但都不尽如人意，因为这种方法是在加入 Cp 后每隔一定时间间隔取一部分反应液，停止反应，用氨基酸分析测定释放出来的游离氨基酸，以释放氨基酸的量为纵轴，作用时间为横轴得一曲线，这样来分析释放顺序，一般约几个，且先后顺序的区别也不是很容易的。

梁宋平教授实验室采用羧肽酶和 MALDI－TOF 联用测定蛋白质的 C 端序列[30]。

作者指出用 C 端序列仪分析重组蛋白质的 C 端很有意义，但也有不少问题，例如 C 端如是 Pro，则完全不能进行，对 Cys、Thr、Ser 的分析很困难，其他氨基酸的 C 端分析则较顺利。在作者实验室中，最好的结果可测定 10 个 C 端氨基酸序列。此法的主要困难在产率、重复效率等不够理想，尚有待进一步完善[31~33]。

第十一节　蛋白质的二硫键分析

二硫键和巯基与蛋白质活性有着密切的关系，早期测定巯基用对氯汞苯甲酸(PMCB)法，后来用 5,5′-二硫-双(2－硝基苯甲酸)[5,5′-dithio bis(2-nitro benzoic acid)，DTNB]和 *N*－乙基顺丁烯二酰亚胺(NEMI)，可以研究巯基分布以及与活性的关系。

遗传工程产品的二硫键是否正确配对，是一个重要的问题，在 IL-2 中，如果形成错配对，不但生物活性只有原来的 1/400，而且在医用后会引起不必要的麻烦。因此分析二硫

键正确配对是一项必不可少的程序。目前有几种方法可以借鉴。

一、不同同位素标记法

人 IL-2 有三个 Cys，由天然 IL-2 的结构分析得知 Cys58 和 Cys105 形成二硫键，而 Cys125 以游离巯基形式存在。为了阐明重组 IL-2 的二硫键是否以正确形式配对，在 pH 8.5下用^{3}H-碘代乙酸处理 IL-2，在这种情况下，仅游离的巯基才能与^{3}H-碘代乙酸作用形成相应的衍生物。然后再还原和用^{14}C-碘代乙酸烷化。因此^{3}H 标记仅在肽 11 上（Cys125），而肽 11 上基本上无^{14}C 的放射性，而^{14}C 的放射性全在肽 12 和肽 14 上，因此，说明 rIL-2 的二硫键结构和天然 IL-2 是一致的[34]。

二、4-乙烯吡啶标记法

超氧化物歧化酶（SOD）也有三个 Cys，由一级结构分析指出 Cys6 为游离巯基，Cys55 和 Cys144 形成二硫键。通常可用同位素标记方法来判别二硫键的位置，如上述方法一样，但应用同位素在国内有其不方便之处，我们采用化学修饰法来判断，用 4-乙烯吡啶与天然的 SOD 作用，此时仅 Cys6 和化学修饰剂形成衍生物，此衍生物在 254nm 有特殊吸收，因此在胰蛋白酶酶解后，在 254nm 下只有一个肽峰，而还原型 SOD 经同样处理，则有三个肽峰，经 Edman 法测三个肽的序列，得知肽 1 含 Cys6，肽 2 含 Cys55，肽 3 含 Cys144[35]（图 1.7）。

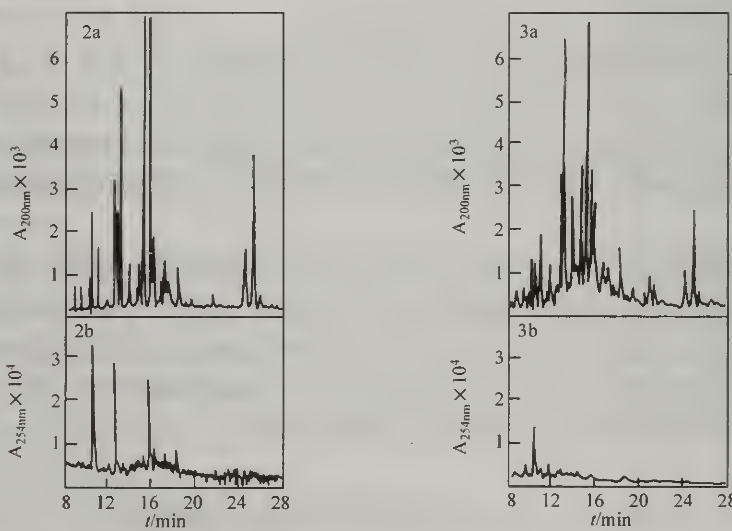


图 1.7　毛细管电泳分析识别二硫键和含半胱氨酸的肽

电泳条件：样品浓度：1μg/μl，进样 3.5～7.0nl；缓冲液：100mmol/L 磷酸缓冲液- 10mmol/L NaCl，pH 2.5；毛细管直径 50 μm，72cm（50cm 至检测器）；电压：20kV，电流：42 μA；检测：200nm（上图）或 254nm（下图）；左图是还原型 SOD，右图是氧化型 SOD

三、直接色谱法

直接分析二硫键的异构物。IL-2 在碱性条件下可以形成二硫键异构体(Cys58 与 Cys125 或 Cys105 和 Cys12 形成二硫键);在反相色谱中,它们的疏水性弱些,较早被洗脱出柱[36]。

四、质　谱　法

近年来,质谱定位二硫键已成为有效的方法。最近,我们以天花粉凝集素为材料,用质谱法测定此凝集素有 4 对分子内二硫键,2 对分子间二硫键。图 1.8 显示 N 端肽的分析,Asn-5 用 PNGase F 酶解除去糖链后转化成 Asp,图 1.8a 是没有还原时肽的质谱图,其双电荷图 m/z 为 1 637.0,指出是分子间的二聚体。图 1.8b 是此肽还原后的串联质谱图,说明 Cys3 和 Cys14 之间有分子内二硫键,且 Cys2 之间的分子间二硫键被 NEM 处理还原[37]。

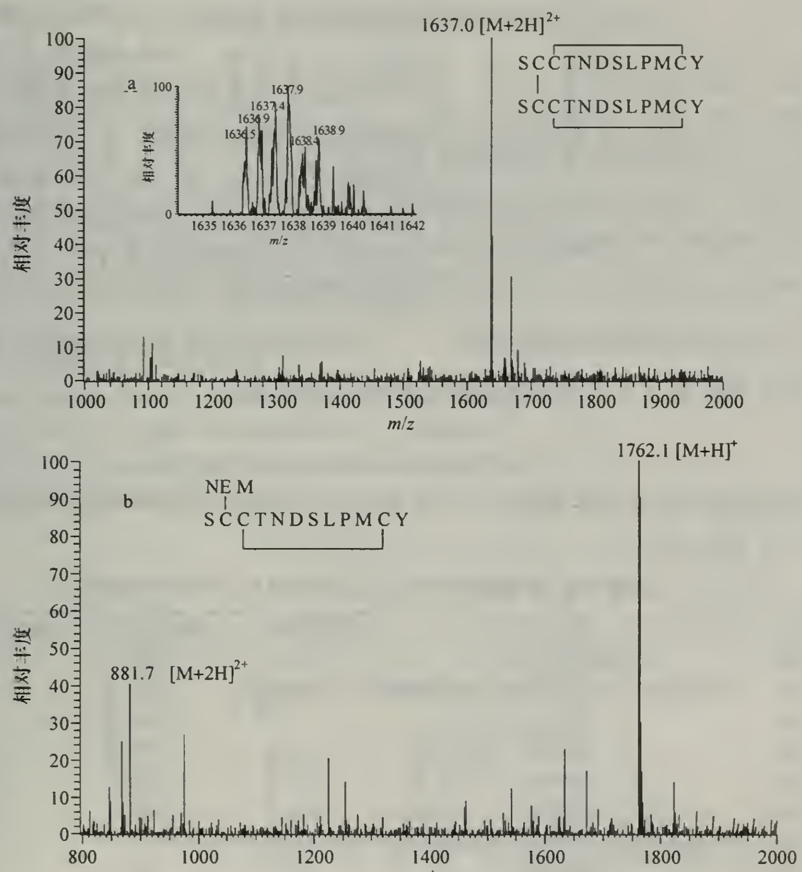


图 1.8　用质谱法定位蛋白质的二硫键

第十二节 突变点的分析

天然人白细胞介素-2(IL-2)由133个氨基酸组成,其中有一个游离半胱氨酸(Cys,位于第125位),第58位和第105位的半胱氨酸形成二硫键,此二硫键为活性所必需,而二硫键的错配对,可降低其生物活性和形成抗原性。用蛋白质工程方法将IL-2的125位Cys改为丝氨酸(Ser)或丙氨酸(Ala),其热稳定性等方面都有较明显的改善。

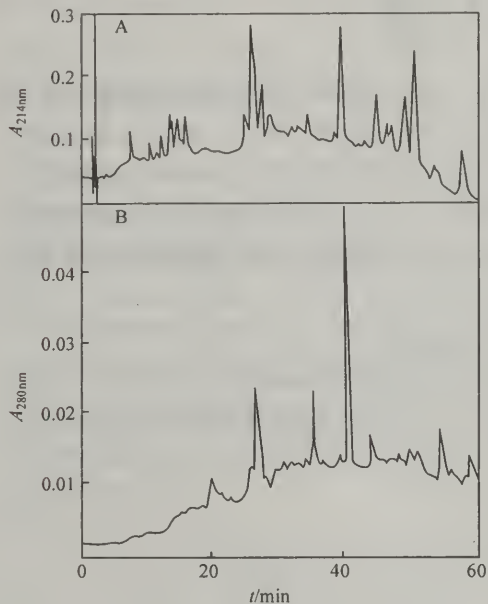


图1.9 IL-2(Ser125)的胰蛋白酶酶解肽谱

为了鉴定新型IL-2的生物工程突变点,我们将IL-2用N-对甲苯磺酰苯丙氨酰氯甲酮-胰蛋白酶(TPCK-胰蛋白酶)水解后,进行了RP-HPLC分析分离。为了搜寻点突变所在肽段,可以用下述方法进行。①天然IL-2和新型IL-2的酶解图谱比较法,寻找有差别的肽进行顺序分析。②荧光标记Cys残基,比较标记前后新型IL-2酶解图谱的差别,从而鉴别出点突变的肽。③用同位素标记法。④由于IL-2只有一个色氨酸残基,而蛋白质的280nm紫外吸收主要是色氨酸,其次是酪氨酸和苯丙氨酸的贡献。所以,我们采用酶解后分别于214nm和280nm处检测肽谱,在280nm肽谱中,光吸收特高的峰就应该是含有色氨酸(Trp)的肽段(残基121~133),可以简便地作出判断(图1.9)。

分离到的肽段再经过毛细管电泳(CE)鉴定确定为均一,最后进行了该肽的氨基酸序列分析,证实为:

121 125 130 133

Trp-Ile-Thr-Phr-Ser-Gln-Ser-Ile-Ile-Ser-Thr-Leu-Thr

从而由蛋白质化学分析方法直接证明了在新型IL-2中Ser125确实取代了天然IL-2中的Cys[24](表1.3,图1.10)。

表1.3 毛细管电泳-质谱联用分析S-125突变的IL-2胰蛋白酶解肽段

肽段	位置	序列	保留时间/min	理论分子质量/Da	实测分子质量/Da
T1	-1~8	MAPTSSSTK	19.48	909.0	908.5
T2	9~35	KTQLQLEHLLLDLQMILNGINNYKNPK	14.71	3192.7	3192.4
	36~38	LTR		388.5	
T3	39~43	MLTFK	15.88	638.8	638.4
T4	44~48	FYMPK	16.37	684.8	684.5
T5	49~54	KATELK	10.37	688.8	688.5
T6	55~76	HLQC* LEEELKPLEEVLNLAQSK	17.55	2621.0	2621.4
T7	77~83	NFHLRPR	7.97	939.1	939.2
T8	84~97	DLISNINVIVLELK	17.97	1582.9	1581.8
T9	98~120	GSETTFMC* EYADETATIVEFLNR	31.97	2684.9	2684.2
T10	121~133	WITFS* QSIISTLT	39.73	1496.7	1496.8

进一步,我们受到 HPLC－二阶微分光谱分析的启发,也建立了这种技术并用于重组蛋白质的 C 端分析[23]。

最近,我们建立了 CE-MS 联用技术用于蛋白质突变点的分析中,将整个分析水平又提高到一个新阶段[38]。

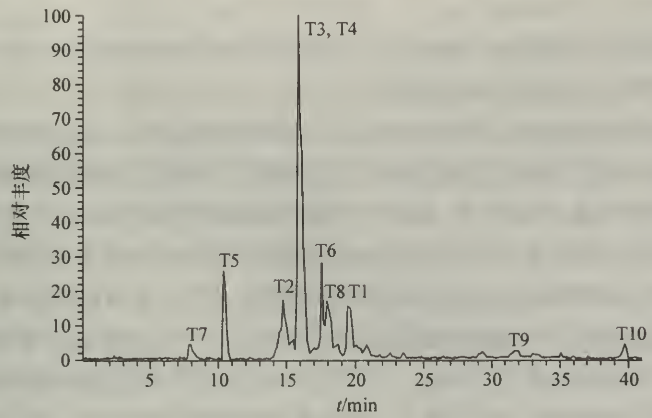


图 1.10 毛细管电泳-质谱联用测定 IL-2 突变

液-质联用已普遍用在重组蛋白的鉴定中。TNF-α 突变株在 N 端和 C 端都有改变,早期我们分别采用 N 端和 C 端分析 Edman 降解法分别测定。近来用液-质联用,分析水平又上了一个层次。TNF-α 用胰蛋白酶酶解,然后进行液-质分析,肽的分子质量与理论值一致,均为正常肽;如与理论值不一致,则可能是突变点所在的肽。TNF-1 理论分子质量是 17 522.6Da,突变株分子质量为 16 864.3Da,实验测得分子质量为 16 864.8Da。肽谱中有两段肽与理论值不一致,经 MS/MS 测得的肽序列,证实确为 N 端和 C 端突变肽(图 1.11)。

MRRRKPVA HVVANPQAE
VRSSSRTPSDKPVA HVVANPQAEG QLQWLNRRAN ALLANGVELR
DNQLVVPSEG LYLIYSQVLF KGQGCPSTHV LLTHTISRIA
VSYQTKVNLL SAIKSPCQRE TPEGAEAKPW YEPIYLGGVF
QLEKGDRLSA EINRPDYLDF AESGQVYFGI IAL
SGQVYFGI IAQ

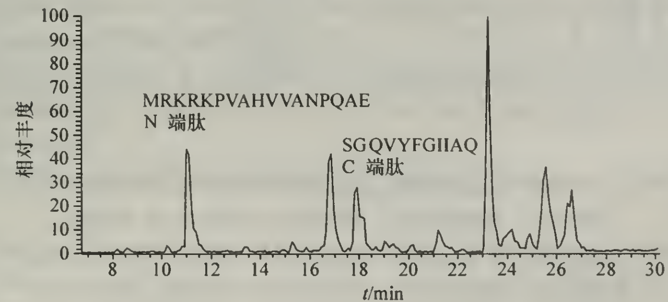


图 1.11 TNF-α 的突变点分析

第十三节　原位分析和快速分析

一、原位分析

原位分析是指蛋白质在 SDS-PAGE 后，电印迹至 PVDF 或 ProBlot 膜上，然后将此膜直接置于蛋白质序列仪上进行序列分析[13,14]。

在遗传工程产品的研究中，原位分析具有非常重要的作用。当重组蛋白质表达后，需确认是否有预定的表达产品及其数量的多少。如果没有原位分析法，则需将粗的表达产物进行工作量很大的分离纯化工作，然后对大量不同的积分进行分析才能得出结论。

目前，人们常在分离之前，进行 SDS-PAGE 分析，根据电泳谱上的条带和分子质量来估计有否预定的产品和产量，这只是一种粗略的估算。进一步的验证需将相应分子质量的条带电印迹到疏水膜上，再置此膜到蛋白质序列仪上分析，如果这一序列和预定表达的蛋白质序列一致，就能得出肯定的结论。可以认为这是非常简便而又有效的方法。

在肯定了有表达产物和一定的量后，必须进行分离纯化的工作，在分离纯化的每一步，常用 SDS-PAGE 来判断和评价每一程序的效果，提纯倍数，活性得率和蛋白质回收如何，等等。

二、快速分析

文献上报道了两种快速的色谱方法：无孔径载体色谱柱（non-porous resin column, NPR）和灌注色谱（perfusion chromatography）[39]。

我们就无孔径载体色谱柱（NPR）在蛋白质和多肽等遗传工程产品中的快速分析作了报道（图 1.12）。通常用经典的 HPLC 柱做一次样品分析约需 30min，而用无孔径载体柱只需 2～5min，这种快速高效的方法对遗传工程产品的分离过程跟踪尤为有利[40]。

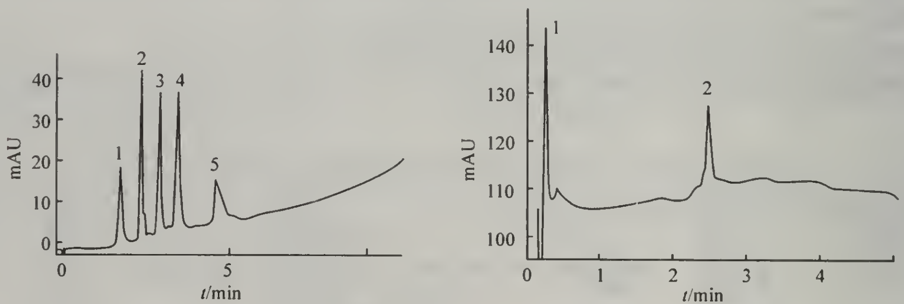


图 1.12　标准蛋白质和遗传工程产品在 NPR-18 柱上的快速分离

左图样品：1. 核糖核酸酶；2. 胰岛素；3. 细胞色素 c；4. 溶菌酶；

5. 牛血清白蛋白（各 1μg）；NPR-18 柱：35mm×4.6mm I.D；检出：220nm。

右图样品：1. rhEGF；2. γ-IFN；流速：1.5ml/min；梯度：TFA-ACN

第十四节　蛋白质翻译后修饰的分析

蛋白质翻译后修饰包括磷酸化、糖基化等。蛋白质一级结构可以从基因序列演绎，但翻译后修饰的信息几乎无法仅从基因序列得到。而众所周知的是，这些翻译后修饰特别是磷酸化、糖基化对于蛋白质的功能乃至细胞的功能都十分重要。因此，这些修饰的鉴定对于重组蛋白质控，蛋白质组特别是功能蛋白质组的研究意义重大。Edman 降解法对于这些修饰的鉴定无能为力，而采用质谱法可通过特征离子监测的方法很快确定磷酸化肽，通过串联质谱确定磷酸化位点[41]。在糖蛋白分析方面，质谱更显示了独特的优势，不仅可以通过质谱、蛋白酶解和糖苷酶酶解相结合的方法寻找糖肽，鉴定糖基化位点，还可依靠 MS/MS 和 MSn 分析糖链组成、结构甚至分支情况。我们采用质谱方法鉴定了一种 C 凝集素的 N 糖链结构和位点[42]；质谱也用于鉴定蛋白质的磷酸化位点；用质谱法我们还发现了一种新型的唾液酸[43]。此外，还用毛细管电泳建立了糖蛋白的糖型分析[44,45]。这些将在有关章节中讨论。

（夏其昌）

参考文献

1 夏其昌.超滤法及其应用.见：鲁子贤主编.蛋白质和酶学研究方法，北京：科学出版社，1989.42～52

2 虞荣华，鲁子贤.蛋白质的盐析法.见：鲁子贤主编.蛋白质和酶学研究方法，北京，科学出版社，1989.17～28

3 吴克佐.酶或蛋白质的结晶.见：鲁子贤主编.蛋白质和酶学研究方法，北京，科学出版社，1989.17～28

4 Bauuer J. Cell Electrophoresis. CRC Press, 1994

5 Song JF, Liu T, Shen X, *et al*. Application of free－flow electrophoresis to the purification of trichosanthin from a crude product of acetone fractional precipitation. Electrophoresis, 1998, 19(7): 1097～1103.

6 Chibnall AC. The total nitrogen content of egg albumin and other protein. Biochem J, 1943, 37: 354～359

7 任梅轩.蛋白质的透析，冷冻干燥及一些基本分析方法.见：潘家秀等编著.蛋白质化学研究技术，北京：科学出版社，1973.1～15

8 Levin P, Brauer RW. The biuret reaction for the determination of proteins: an improved reagent and its application. J Lab Clin Med, 1951, 38: 474

9 Lowry OH *et al*. Protein measurement with the Folin phenol reagent. J. Biol. Chem., 1951, 193: 265

10 Smith PK *et al*. Measurement of protein using bicinchoninic acid. Anal Biochem, 1985, 150: 76～85

11 Bradford MM. A rapid and sensitive method for the quantitation of microgram quantities of proteins utilizing the principle of protein dye binding. Anal Biochem, 1976, 72: 248～254

12 吴如丹.溶液中蛋白质浓度的测定.见：鲁子贤主编.蛋白质和酶学研究方法，北京：科学出版社，1989.1～7

13 Matsudaira P. Sequence from picomole quantities of proteins electroblotted onto polyvinylidene difluoride membrane, J Biol Chem, 1987, 262: 1003

14 Yuen SW, Chui AH, Wilson KJ *et al*. Microanalysis of SDS－PAGE electroblotted proteins. Biotechniques. 1989, 7(1): 74～83

15 Stone KL, Willams KR. Enzymatic digestion of proteins and HPLC peptide isolation. In: Matsudaira P ed. A pratical guide to protein and peptide purification for microsequencing. Academic Press, Inc. 1993. 43～69

16 Mann KG, Fish WW. Protein polypeptide chain molecular weight by gel chromatography in guanidinium chloride. Methods Enzymol, 1972, 26: 28～42

17 Nielsen TB, Reynolds JA.. Measurement of molecular weight by gel electrophoresis. Methods Enzymol, 1978, 48: 3～10

18 黄培勇等.重组人表皮生长因子的纯化和鉴定（内部材料）.
19 屠红旻，夏其昌.毛细管电泳及其在生命科学中的应用.生命的化学，1994，14(6)：36～39
20 夏其昌，徐松琴，蔡嘉坤等.应用乙醇系统的高效液相色层(HPLC)可在 10 微微克分子水平定量分析氨基酸.生物化学与生物物理学报，1985，17：747
21 夏其昌，吴高德.反相色层法微量分析 PTC-氨基酸.生物化学与生物物理学报，1989，21：465
22 王海波，徐来根，吴高德等.人 γ 基因工程干扰素 C 端的结构分析及功能作用.生物工程学报，1992，10：50～55
23 张丽华等.HPLC 在线二阶微分光谱法检测基因工程产品的 C 端肽.科学通报，1995，40：1314
24 吴高德等.新型白细胞介素-2 突变点的分析鉴定.生物工程学报，1992，8：314～317
25 Sanger F. The free amino groups of insulin. Biochem J, 1945, 39: 507
26 Edman P. Method for detemination of the amino acid sequence in peptides. Acta Chem. Scard, 1950, 4: 283
27 Wittmann-Liebold B, Hirano H, Kimura M. Manual microsequence determination of proteins and peptides with the DABITC/PITC method. In: Wittnann-Liebold B *et al* eds. Advanced methods in protein microsequence analysis Springer-Verlag, 1986
28 金善炜.蛋白质顺序测定技术.有机化学，1982，5：384～386
29 Niu CF, Fraenhel-Conrat H. C-terminal amino acids and peptides by hydrazinolysis. J Am Chem Soc, 1955, 77: 5882
30 陈平，梁宋平.一种优化的采用 MALDI-TOF 分析多肽 C 端序列的方法，中国生物化学与分子生物学报，2001，17：670～676
31 Hawke, DH *et al*. Studies on C-terminal analysis. In: Hugli TE ed Technique in protein chemistry. Academic Press Inc., 1989, 59
32 Inglis AS *et al*. Methods in Protein Sequence Analysis. Birkhansen Verlag, 1990. 23
33 屠红旻，夏其昌.蛋白质和多肽 C-端氨基酸顺序研究进展.生物化学与生物物理进展，1993，20：181
34 Jones BN, Stein S. Fundamentals of protein bio-technology. Marcel Dekken Inc., 1990. 181
35 屠红旻，夏其昌.毛细管电泳测定牛红细胞超氧化物歧化酶的巯基位点.生物化学杂志，1992，8：570～573
36 Browing JL. *et al*. Disulfide scrambling of interleukin-2: HPLC resolution of the three possible isomers. Anal Biochem, 1986, 155: 123～128
37 Zeng R, Xu Q, Shao XX *et al*. Determination of the disulfide bond pattern of novel C-lectin from snake venom by mass spectrometry. Rapid Commun Mass Spectr, 2001, 15: 2213～2220
38 Liu T, Shao XX, Zeng R *et al*. Analysis of recombinant and modified proteins by capillary zone electrophoresis/electrospray ionization-tradem mass spectrometry. J Chomatography A, 1999, 855: 695～707
39 Regnier F. Perfusion chromatography. Nature, 1991, 350: 634～635
40 张丽华等.HPLC 无孔径载体柱用于快速分离蛋白质和多肽.高技术通讯，1995，5(2)：45～47
41 Che FY, Shao XX, Xia QC. Identification of phosphorylation sites of proteins by high performance liquid chromatography-electrospray ionization-quadrupole ion trap mass spectrometry. Sci Chin (Series C), 2000, 43: 561～563
42 Zeng R, Xu X, Shao XX *et al*. Characterization and analysis of a noval glycoprotein from snake venom using liquid chromatography-electrospray mass spectrometry and Edman degradation. Eur J Biochem, 1999, 266: 352～358
43 Che FY, Wang KY, Xia QC. Determination of sialic acid content in glycoproteins by capillary zone electrophoresis. Electrophoresis, 1999, 20: 2930～2937
44 曾嵘，邵晓霞，王克夷等.液相色谱-电喷雾质谱法分析牛胰核糖核酸酶 B 酶解肽谱，糖基化位点及糖型.生物化学与生物物理学报，1999，31(6)：695～700
45 Che FY, Song JF, Shao XX *et al*. Comparative study on the distribution of ovalbumin glycoforms by capillary electrophoresis. J Chromatogrphy A, 1999, 849: 599～608

第二章 蛋白质结构分析

第一节 蛋白质样品的准备

进行结构分析的蛋白质样品，一般要求纯度>95%。这样的样品，通常是用色谱和电泳精制过的。

一、SDS-PAGE 纯化的样品

从 SDS-PAGE 上分离的蛋白质必需经过特殊处理，使 SDS 浓度低于 0.0005%，否则 SDS 会影响到胰蛋白酶的酶解，也会大大降低 HPLC 分肽时的分辨率[1]。

有报道指出，在 2mol/L 尿素存在下，用胰蛋白酶消化蛋白质时，如果 SDS 浓度>0.05%，则完全阻止了胰蛋白酶和内肽酶 Lys-C 的作用。

虽然有报道用 HPLC 或用盐酸胍沉淀除去 SDS，但不易掌握，尤其是共价结合到蛋白质分子上的 SDS 分子更难除去。

在 SDS-PAGE 后，采用电印迹(blotting)到 PVDF 膜上，也是常用的方法[2]，因为 PVDF 膜结合蛋白质能力的专一性远远超过其结合盐、氨基酸和去垢剂，因此，蛋白质能和这些"污染物"分开。经过这样处理后，能用于 BrCN 化学裂解或酶解[3]。

二、HPLC 纯化的样品

如果样品是用反相 HPLC 纯化的，因为流动相是挥发性溶剂，它们在真空离心干燥或冻干样品时除去；如果蛋白质是用离子交换 HPLC 或疏水 HPLC 或凝胶过滤 HPLC 精制的，则样品要进行脱盐。

脱盐或稀蛋白溶液的浓缩可用 0.1% 三氟乙酸(TFA)-乙腈(ACN)系统的短柱反相 HPLC；也可采用一次性的 Sek-Pak，因为蛋白质是水溶性的，所以先用能与水互溶的甲醇或乙腈(2ml)润之，再用 5ml 纯水洗涤；然后样品液体通过 Sep-Pak 柱时，蛋白质保留在柱上，小分子的盐类流过柱体，用水溶液充分洗涤后，最后用甲醇或乙腈抽提蛋白质。

小的超滤装置对浓缩蛋白质也是有效的[4]。此外，有机溶剂或 TCA 沉淀也可以考虑。

三、蛋白质样品中的二硫键和巯基(—SH)

当蛋白质样品纯度达到要求，盐和小分子也除得比较"干净"，在进行酶解前，还要转变半胱氨酸残基成其他的稳定的衍生物。因为半胱氨酸的巯基在分肽过程中易自动氧化成各种产物，包括形成无序的二硫键，因此需要将半胱氨酸残基先转变成稳定的衍生物。

常用有碘代乙酸烷化，因为这个反应是快速和专一的；同时这类试剂保存在暗处是很

稳定的。而且，这类试剂的矫作物(artifact)容易得到放射性标记物。蛋白质或肽在导入带电荷基团后也增加了其在水溶液中的溶解度。

胱氨酸也可在还原后再修饰。

1. 还原和羧甲基化

将 1～20mg 蛋白质溶解在含 6 mol/L 盐酸胍的 0.1mol/L Tris-1mmol/L EDTA 的缓冲液中，pH8.3，加 DTT 至 2 mmol/L 或再多一些(过量于存在于蛋白质中的二硫键)，通氮气，然后迅速封闭，37℃反应 1h，再加入 50mmol/L 的碘代乙酸(已用 NaOH 中和)，其量为 1.1 倍(分子比)过量于溶液中的巯基数(蛋白质中的—SH 和 DTT 中的—SH 总和)，典型为 4.5～5 mmol/L，再次充氮气，于暗处 37℃保温 1h。加 2-巯基乙醇至最后浓度 1%(*V/V*)结束反应，然后对 50 mmol/L NH_4HCO_3，含 0.01% 硫二甘醇(thiodiglycol)充分透析，冻干。

注意：碘代乙酸必须是无色的，如果有碘存在，呈淡黄色，会引起巯基迅速氧化，阻碍了烷化反应，而且可能有副反应，使色氨酸修饰。

碘代乙酰胺(iodoacetamide)也常代替碘代乙酸用于半胱氨酸残基的烷化。这常用于避免半胱氨酸被导入负电荷，而用中性的碘代乙酰胺。另一种情况是变性剂不用盐酸胍，而用 SDS，这时如果用碘代乙酸，反应往往进行得很慢，而且不完全，这是因为碘代乙酸的负电荷和蛋白质-SDS 胶束(微团)上的负电荷相互排斥。

2. 还原和吡啶乙烯化

用 4-乙烯吡啶(4-vinylpyridine)处理同上，通氮气保温过夜[5]。

第二节　蛋白质的化学裂解

一、裂解于甲硫氨酸(Met)

溴化氰(BrCN)是首选的化学试剂，专一裂解在甲硫氨酸的羧基端，对多数蛋白质裂解效率为 90%以上[6]。BrCN 在酸性条件下，裂解在甲硫氨酸的 C 端肽键上，将甲硫氨酸转化成高丝氨酸(Ⅱ)(homoserine)和高丝氨酸内酯(Ⅰ)(homoserinelactone)的混合物，它们之间也能相互转化，如图 2.1 所示。

对 Met-Thr、Met-Ser 键，则常常是低产率的，可能是β羟基产生下面的一系列反应(Ⅲ和Ⅳ)，最后形成高丝氨酸残基(Ⅴ)，而未裂解此肽键，如图 2.2 所示。Met-Cys 键也不易裂解。

典型的反应常常在 20℃、70%甲酸(*V/V*)中进行 24h，副反应是部分 Asp-Pro 键会断裂，一些酰胺基团会部分丢失，部分 Asn-Gly 会产生重排或环化，有时 Trp 会氧化。

操作：蛋白质先用 5%巯基乙醇在 pH 8.0 室温下处理 24h，加色胺(tryptamine)或色氨酸(以每分子甲硫氨酸加 5 分子色胺或色氨酸计算)以防止色氨酸氧化。

蛋白质以 1～5mg/ml 浓度溶解在 70%甲酸中，加入 50 倍过量，小体积的 BrCN/70%甲酸，通氮并置于暗处，室温放置 16～24h，结束后加 15 倍体积的水，冻干，重复加水冻干。

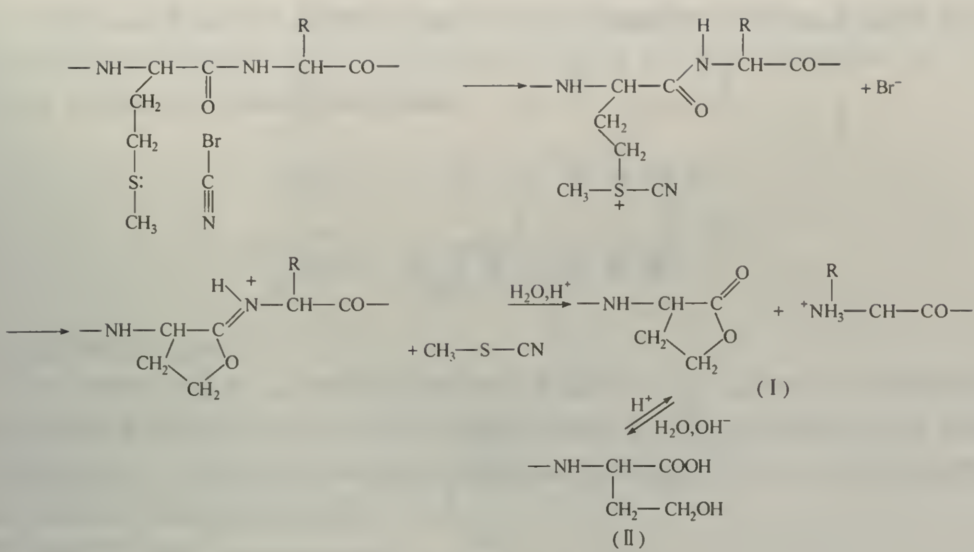


图 2.1　溴化氰裂解甲硫氨酸的 C 端肽键

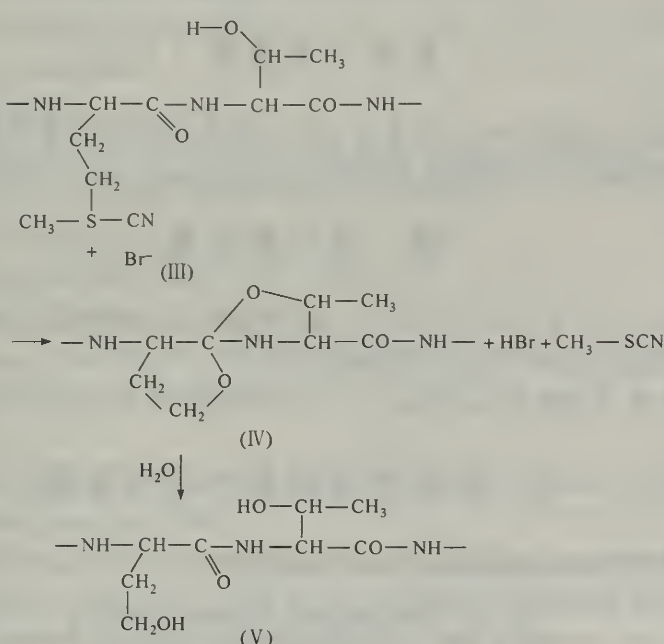


图 2.2　溴化氰裂解 Met-Thr、Met-Ser 键

二、裂解于半胱氨酸(Cys)

半胱氨酸能和 5,5′-二硫-双(2-硝基苯甲酸)(DTNB)(Ⅰ)在 pH 8.0 下形成 S-cyanocysteine残基,过量试剂和副产物在酸性条件下用凝胶过滤法除去,而 S-cyanylted 的蛋白质(Ⅱ)在 pH 9.0 下保温使肽键断开,由 S-cyanocysteine 形成氨端 iminothiozolidine-4-arboxyl 残基(Ⅲ)和释放前一残基的羧端(Ⅳ),由于(Ⅲ)已环化,不能再用 PITC 法自动降解,但可进

一步分肽和做氨基酸分析，或用稀酸水解在天冬氨酸上再进行分析[7]，如图 2.3 所示。

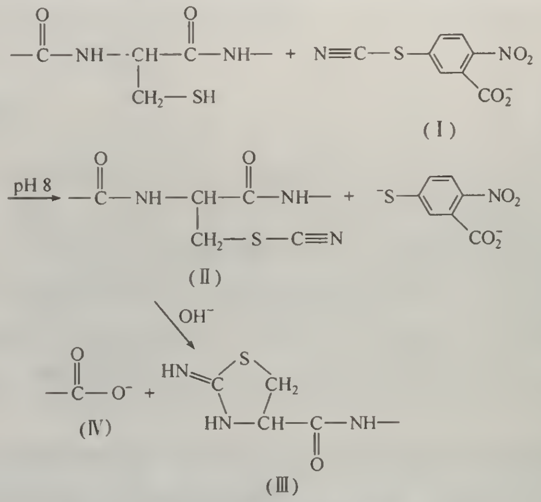


图 2.3　裂解于半胱氨酸

三、裂解于色氨酸（Trp）

常用试剂有 BNPS-Skatole，效率为 15%～60%，副反应是酪氨酸，组氨酸也能作用[8,9]。还有报道用 *O*-亚碘酰基苯甲酸（*O*-iodosobenzoic acid）[10]和 DMSO/HCl/HBr。

四、部分酸水解

用 0.03mol/L HCl 于 105℃ 反应 6～18h，裂解在 Asp 的羧端。0.0125mol/L HCl（pH 2）真空 108℃ 反应 2h，60% Asp-X 键裂解[11]。

温和酸水解，肽键断在 Asp-Pro。

五、选择性及限制性化学裂解

选择性化学裂解如羟胺裂解在 Asn-Gly 键上[12]。

表 2.1　常用的酶解和化学裂解的试剂和方法

裂解试剂	专一性	文献	裂解试剂	专一性	文献
内肽酶 Lys-C	Lys-X	[13]	胰凝乳蛋白酶	Trp-X，Tyr-X	
Achromobacter 蛋白酶 I	Lys-X	[14]		Phe-X，Leu-X	
内肽酶 Arg-C	Arg-X	[15]		Met-X	
内肽酶 Asp-N	X-Asp		溴化氢(BrCN)	Met-X	[20]
	x-cysteic acid	[16]	BNPS-Skatole*	Trp-X	[21]
Asn 内肽酶	Asn-X	[17]	Iodosobenzoic acid	Trp-X	[22,23]
V8 蛋白酶	Glu-X		N-氯代琥珀酰亚胺	Trp-X	[24]
	Asp-X	[18]	2-硝基-5-硫氢基苯甲酸	Cys-X	
胰蛋白酶	Lys-X				
	Arg-X	[19]			

* 2-(2-硝基苯亚磺酰胺)-3-甲基-3-溴代假吲哚。

德国 Stahl 用 BrCN 和蛋白质作用,在不同的时间间隔取样,得到一系列裂解在蛋白质不同部位的甲硫氨酸产物,并进行质谱分析,这有助于蛋白质全序列的演绎。表 2.1 提供了常见的酶解和化学裂解的试剂和方法。

第三节　蛋白水解酶酶解

一、常用蛋白水解酶

(1) 胰蛋白酶 (trypsin)

专一作用在 Arg 和 Lys(碱性氨基酸)的羧基端,但对 Arg-Pro 和 Lys-Pro 键没作用。我们常将胰蛋白酶配成 1mg/ml,贮存在 1mmol/L HCl 中,－20℃保存数月,而未观察到酶的活性损失。如果样品溶解度差,可适量加入尿素增加溶解度,因为胰蛋白酶在 2mol/L尿素中,仍具有充分的活性。

(2) 胰凝乳蛋白酶 (α-chymotrypsin)

专一作用在芳香族氨基酸的羧基端。

(3) 内肽酶 Lys-C (endoproteinase Lys-c)

专一作用在 Lys 的羧基端。

(4) V8 蛋白酶 (staphylococcal protease)

专一作用在 Asp 和 Glu(酸性氨基酸)的羧基端。pH 8.0 的磷酸缓冲液时,作用在 Glu-X 和 Asp-X;pH 4.0 的乙酸缓冲液时,仅作用于 Glu-x 键。

(5) 梭菌蛋白酶 (clostripain)

专一作用在 Arg 的羧基端[25]。

(6) 凝血酶 (thrombin)

专一裂解在 Arg－Gly 键上[26],在纤维蛋白原工作中,很大程度上还依赖于邻近氨基酸序列。近来还常用作融合蛋白的裂解点[27]。

(7) 其他

还有专一性较广的胃蛋白酶(pepsin)、枯草杆菌蛋白酶(subtilisin)、木瓜蛋白酶(papain)、嗜热菌蛋白酶(thermolysin)等。

应该指出胰蛋白酶、胰凝乳蛋白酶均从胰脏制备,常常不易完全分开,而是你中有我,我中有你,专一性不佳,易造成混乱。应采购 TPCK－胰蛋白酶(抑制了胰凝乳蛋白酶活性)和 TLCK－胰凝乳蛋白酶(抑制了胰蛋白酶活性),才能保证酶的专一性。

二、酶解操作条件

将 30μl(100μg)蛋白质,100mmol/L NH_4HCO_3,pH 8.5,5mmol/L DTT,50℃作用 15min,在冷至室温后,加 5μl 100mmol/L 碘代乙酸,室温 15min,加 60μl 水,最后加酶 5μl [蛋白样品:酶＝25:1 (质量比)],37℃保温 20～24h,冻干样品或直接注射样品(用稀 TFA 酸化至约 pH 2.5)到反相 HPLC 柱中分析。

三、讨　　论

(1) 蛋白质的量和浓度

酶解不能正常进行的主要原因之一是蛋白质的量不足。精确定量蛋白质是必需的，一般用氨基酸分析来定量微量的蛋白质。

蛋白质浓度通常要求 500μg/ml，最少不得低于 100μg/ml，否则酶的消化难以进行，体积在 50～100μl。

(2) 缓冲液、盐和去垢剂

酶作用可以被缓冲液中的各种因子所抑制或完全阻止。在标准的 2mol/L 尿素，0.1mol/L NH_4HCO_3 系统中，SDS 含量为 0.0005% 时，就会影响酶的消化。0.005% SDS 会减少许多酶解肽段的产量。类似的，染料考马斯兰和两性电介质也会阻止酶的消化。高浓度的盐(1mol/L NaCl)也会影响蛋白水解酶的活性；2mol/L 以上浓度的尿素会影响羧甲基化和蛋白酶的消化。

(3) 羧甲基化

Stone 等报道，转铁蛋白用酶解时，羧甲基化的转铁蛋白酶解很好，而没有羧甲基化的转铁蛋白，酶解程度很差。

(4) 蛋白酶的量

酶和蛋白质的质量比为 1∶25，如果分子质量大的蛋白质(＞6 万 Da)，则比例要减低至1∶50。胰蛋白酶要用序列级的商品。

(5) 原位酶解

将 SDS-PAGE 上的蛋白质样品从凝胶上解析下来是很困难的，文献报道有很多方法，但实用性差。目前流行的方法是将蛋白质电印迹(blotting)至 PVDF 类的膜上，直接进行序列分析或直接在膜上进行酶解，后者称为原位分析(*in situ*)。

操作如下：膜片剪成 1mm 细条，放在 Eppendrof 管中，加 1.2ml 0.5%(*m*/*V*) PVP-40/0.1mol/L HAc，37℃ 保温 30min，PVP-40 的作用是阻止在酶解时蛋白酶吸附到硝酸纤维素膜上，除去 PVP-40 后，用水充分洗涤膜至少 5 次以洗净残留的 PVP-40(因为 PVP-40 有很强的紫外吸收，在 RP-HPLC 中，PVP-40 于 30%～40% 有机溶剂浓度时洗下来，它在 215～220nm 有很高的吸收，在 260～289nm 吸收很小)，所以在 HPLC 前需完全除去。膜片再用酶解时的缓冲液漂洗，然后进行酶解，一般蛋白酶在 100mmol/L NH_4HCO_3∶ACN＝95∶5，pH8.2，37℃ 酶解过夜，对 V8 蛋白酶则在 100mmol/L 磷酸钠∶ACN＝95∶5，pH 7.8，室温过夜。酶与底物之比为 1∶10 至 1∶20(质量比)；对亚微克级样品，酶与底物之比高达 2∶1。酶解后 －20℃ 保存或立刻用数微升 10% TFA 酸化，上样到 HPLC 柱上分析。

(6) 限制性酶解

有助于蛋白质测序时的序列演绎。最著名的例子是核糖核酸酶(RNase)，它被枯草杆菌蛋白水解酶在 pH 8.0 下作用，活性没有变化，但出现一个新的 N 端丝氨酸。用 Amberlite IRC-50 离子交换柱，0.2mol/L pH6.25 的磷酸钠缓冲液可分离到一个大片段(S 蛋白)和一个小片段(S 肽)。也可用三氯乙酸(TCA)沉淀法，沉淀部分是 S 蛋白，上层

液部分是 S 肽[28]。同法亦可用在糖原磷酸化酶等的研究中。

第四节 蛋白质中存在着封闭的 N 端

有些蛋白质的 N 端是封闭的，尤其在 Ehrlich Ascites 腹水细胞中，80%的蛋白质的 N 端是乙酰化的，因此用 Edman 法不能直接测定其序列，要用一系列的化学或酶处理[29,30]。

如果在蛋白质/肽序列仪上测定不到序列，蛋白质的量肯定又是足够的，则可在仪器上进一步用含水 TFA(三氟乙酸)处理样品，使蛋白质酸水解成许多肽段后再测序，这时会有很多 PTH-氨基酸出现，以此反证此蛋白质的 N 端是封闭着的，再用下述方法除去封闭着的 N 端。

首先，用温和的酸处理可以除去甲酰化的甲硫氨酸(f-Met)，也可能引起丝氨酸、苏氨酸的乙酰化 N—O 重排反应。如果酸处理后还是测不到序列，则进一步用焦谷氨酸氨肽酶除去蛋白质 N 端上的焦谷酰基；假设这一方法仍不见效，再用较强的酸处理或其他酶处理。虽然有乙酰氨基酸释放酶(acylamino acid-releasing enzyme，AARE)，但它不能作用长度在 20 个氨基酸残基以上的蛋白质或肽。因此，封闭的蛋白质首先裂解成小肽，再用 AARE 处理除去 N 端上的乙酰化氨基酸转变成 n-1 肽，然后 n-1 肽的序列再用 Edman 法测定。

一、除去蛋白质中封闭的 *N*-乙酰丝氨酸和 *N*-乙酰苏氨酸

样品置于 1.5ml 的聚丙烯离心管中，加无水 TFA，密闭，40℃保温 1h，然后打开管盖，让 TFA 挥发，干燥，再进行序列分析。

二、除去 *N*-甲酰甲硫氨酸

样品置于 1.5ml 聚丙烯离心管中，加 30μl 0.6mol/L HCl，密闭，25℃水解 24h，打开管盖，真空除去 HCl，然后进行序列分析。

以上两种温和酸水解除去甲酰或乙酰氨基酸，这与时间、温度、酶浓度有很大关系，过分的酸水解会引起蛋白质中肽键的断裂，如对酸不稳定的 Asp-Pro 键；相反，条件不够，则封闭的 N 端基团水解不完全，导致测序的起始点不均一。

三、除去N端的焦谷氨酸残基(pyroglutamate)

样品在序列仪上测不出 N 端，则将载有样品的膜片转移至 1.5ml 聚丙烯微量离心管中，用 0.5%PVP-40 洗膜片(见上节原位酶解部分)，再加 100μl 焦谷氨酸氨肽酶(5μg 于 50mmol/L 磷酸钠 pH7.0，含 10mmol/L DTT) 37℃酶解 5～10h，用水洗膜，干燥、放回序列仪分析。

四、除去其他的 *N*-乙酰氨基酸[31]

五、选择性纯化N端封闭的肽

蛋白质用胃蛋白酶(pepsin)或嗜热菌蛋白酶(thermolysin)作用后(但不能用胰蛋白酶!),酸化到 pH 2,上强酸性阳离子交换柱(例 Dowex 50mm×2mm,H^+ 型)。所有含游离氨基或其他碱性基团的肽都结合至柱上,仅缺少碱性基团的肽通过柱泄出。如果 N 端是封闭着的(没有游离的 N 端氨基),同时因用专一性很广的酶作用,肽较小又不含组氨酸、赖氨酸、精氨酸,则可在排出液中发现 N 端封闭的肽。

此技术的局限性是:大肽(甚至含有碱性基团)可能不结合到树脂上,因为它们不渗入树脂孔径内;相反,一些疏水的肽,特别含有色氨酸残基,即是它们缺少游离的 N 端氨基,也可能结合到柱上。此外,如果存在碱性基团非常靠近 N 端,胃蛋白酶作用不能除去,则该技术也不见效。

应该指出,N 端为谷氨酸的肽在极端 pH 下会环化而转变成封闭,需加以小心。

近年来,由于质谱(MS)在生物样品检测中有了长足的发展,可用 MS 精确测定分子质量的方法来判断封闭的 N 端为何种基团[32]。

第五节　蛋白质中的C端残基的测定

一、羧肽酶法

羧肽酶是最常用的方法,对一些小肽而言,C 端有时最容易测定的,如胰蛋白酶酶解肽段时总是以 Lys 或 Arg 为 C 端;BrCN 裂解的肽总是以高丝氨酸(homoserine)为 C 端;V8 蛋白酶酶解的肽总以 Glu 或 Asp 为 C 端。虽然有非专一的情况,但毕竟出现的概率很少。

羧肽酶酶解是从 C 端开始,一个一个地从 C 端释放出氨基酸残基,这是个动态的释放过程。有的氨基酸释放快些,有的氨基酸释放慢些。对小肽的 C 端序列测定问题不大,但对长肽的序列测定,由于对不同氨基酸残基释放的速度不同,对后面一些氨基酸的序列判断,可能会出现错误。在这方面不如化学降解法。

C 端序列分析是在酶解的不同时间间隔,取出一定量的反应物停止反应后,用氨基酸分析定量氨基酸来排列出其序列,近来用质谱法测定 C 端序列。参见有关章节。

(1) 羧肽酶 A (Cp A)

此酶释放非极性氨基酸速度快,释放 His、Thr、homoserine 也很快,释放 Asp、Ser、Met sulphone 则缓慢,释放 Lys 也较慢,但在高 pH 下则速度加快点;酸性氨基酸的释放也较慢,对 Pro 和 Arg 基本上不起作用。倒数第二个氨基酸残基影响水解速度,但对这一现象作一归纳或解释是很困难的。

此外,如何保证 Cp A 不污染有内肽酶也是一个很重要的问题,通常用二异丙基磷酰氟(Diisopropylfluorophosphate,DFP)处理以抑制丝氨酸蛋白酶的活性。

Cp A 在低离子强度下是不溶性的，通常以悬液形式存放于 4℃。使用前将酶悬浮在水中，离心除去上层液，上层液中可能含有游离的氨基酸，而沉淀部分溶于少量 2mol/L NH_4HCO_3 中，此 Cp A 加到变性的底物蛋白质中（10～100nmol, pH 8.5, 0.2mol/L *N*-ethylnorpholine HAc），酶和底物的分子比为 1∶100，如果底物在此缓冲液中不溶解，则可在 1%SDS（*m*/*V*）或 6mol/L 尿素中进行。有时加正亮氨酸（norleucine, 1mol/mol 蛋白质）为内标。溶液在 25℃保温，隔 0.5, 1, 4, 16h，取一定量反应液（1～5nmol），加等体积 1mol/L HAc 停止酶作用，离心除去蛋白质，上层液真空干燥，然后残余物溶在氨基酸样品缓冲液中进行氨基酸分析。以正亮氨酸作校正，同时有相同程序的空白对照。

如果在酶作用短时间后就释放出大量的氨基酸，则改用低的 Cp A 浓度和作用短时间（5, 10, 20, 60min）取样；相反，如果只有少量氨基酸释放，则可加大酶/底物的比例，或 37℃作用更长的时间。

(2) 羧肽酶 B (Cp B)

从蛋白质/肽的 C 端释放 Lys 和 Arg，Cp B 存放在 −20℃时，非常稳定。

(3) 羧肽酶 Y (Cp Y)

目前 Cp Y 有更广泛的专一性和通用性，一般讲它作用底物 C 端的疏水性氨基酸快些，作用于带电荷的氨基酸慢些。此酶很稳定，某些金属离子（例如 Cu^{2+}, Hg^{2+}）抑制酶活性。Cp Y 制剂中可能污染有酵母蛋白酶 A，这时可加胃蛋白酶抑制剂（pepstatin）抑制，也可用 EDTA 抑制。Cp Y 用量一般为底物的 1/50～1/100，在 25℃作用，于不同时间取反应液酸化后分析游离的氨基酸。

(4) 羧肽酶 P (Cp P)

Cp P 是种酸性羧肽酶，它贮存在 pH 5.5 的 50mmol/L 吡啶醋酸缓冲液中，于 −20℃, 6 个月内非常稳定。其作用于 pH2.5 的 100mmol/L 吡啶甲酸缓冲液中。

二、化学降解法

早年的化学降解法只能测定一个 C 端氨基酸，远远不能满足需求。而羧肽酶法虽能测到 3～5 个序列，但这是一个动态演绎的方法，有可能错判。十多年前，蛋白质化学家就开始用类似 Edman 降解法从 C 端逐个测定氨基酸序列，每次国际蛋白质会议都有报道，直到 1995 年才有商品仪器供应。目前 C 端序列仪一般只能测定 5 个氨基酸残基左右（这和蛋白质的性质有关，有些可测到 10 个左右），在 nmol 水平上；这与 N 端序列还有很大的差距（目前 N 端序列仪灵敏度在 10pmol，可测到 50 个循环）

第六节　蛋白质中一些特殊肽段的检出

一、含巯基肽的分离纯化和分析

巯基在蛋白质性质、结构和功能中具有重要作用，因此含巯基肽段的鉴定是蛋白质化学不可缺少的部分。文献上含 Cys 的肽常用同位素标记的碘代乙酸处理，羧甲基化，然后在 HPLC 肽谱分析时跟踪有标记的肽段。鉴于用同位素在国内不很方便，我们用 4-乙烯

吡啶标记蛋白质中的半胱氨酸,酶解后的肽谱,除用 214nm 检出外,同时还用 254nm 检出,只有被 4-乙烯吡啶修饰的肽在 254nm 才有吸收,所以能方便地检出[33]。

Steinman[34,35]等用碘代乙酸修饰牛红细胞超氧化物歧化酶(SOD),测定了其一级结构和二硫键的定位,确定此酶含有一对二硫键和一个游离的巯基,因此 SOD 可以成为研究巯基修饰的良好材料。碘代乙酸修饰的蛋白质在进行氨基酸序列分析时,X 甲基半胱氨酸的 PTH-衍生物(PTH-CMC)和 PTH-Glu 较难分开,影响序列测定,同时需用同位素标记才能分离到含有巯基的肽段。我们现用 4-巯基吡啶修饰半胱氨酸,可以在分析 PTH-氨基酸时清楚地定位乙基吡啶化半胱氨酸的 PTH-衍生物(PTH-PEC),同时利用 4-乙烯吡啶标记后在 254nm 处有特殊吸收这一性质[33],分别标记还原的 SOD 和天然的 SOD,从而制备含有巯基的肽和定位二硫键。

(一)试剂和仪器

牛红细胞超氧化物歧化酶:Sigma 公司产品。

巯基乙醇,HPLC 级:Fluka 公司产品。

4-乙烯吡啶:Serva 公司产品,重蒸并充氮气,-20℃保存。

盐酸胍:Schwarz/Mann Biotech 公司产品。

TPCK-胰蛋白酶:本实验室自制。

270A 毛细管电泳:Applied Biosystems 公司产品。

150 HPLC 和 RP-300 柱(Aquapore, 30mm × 4.6mm I. D. C-8), Spheri-5(50mm × 1.0mm I. D. ,C-18):Applied Biosystems 公司产品。

477A/120A 蛋白质序列仪:Applied Biosystems 公司产品。

(二)方　　法

(1) 总巯基修饰(还原型 SOD)

300μgSOD 溶于 100μl 6mol/L 盐酸胍-0.25mol/L Tris-HCl-1mmol/L EDTA, pH 8.5 缓冲液中,充入氮气,加入 5μl 10%巯基乙醇水溶液,室温避光反应 2h,然后加入 4μl 4-乙烯吡啶,室温避光反应 1.5h。立即用反相 HPLC RP-300 柱脱盐分离,收集样品后抽干。

(2) 游离巯基修饰(天然 SOD)

300μg SOD 溶于同上的胍-Tris-EDTA 缓冲液中,充入氮气,加入 15μl 4-乙烯吡啶,室温避光反映反应 3h,立即经 HPLC RP—300 柱脱盐分离,收集样品并浓缩抽干。

(3) 胰蛋白酶水解

修饰后的样品溶于 300μl 0.1mol/L NH_4HCO_3 溶液中,加入 3μgTPCK-胰蛋白酶(底物与酶质量比约为 100:1),37℃,避光酶解 4h,再追加同量的酶,反应过夜。10%TFA 调至 pH 3.0,终止反应。

(4) 含巯基被修饰肽段的氨基酸序列分析

调节序列仪中 120A HPLC 分离 PTH 氨基酸洗脱液 A(5%四氢呋喃)中 HAc-NaAc 的离子强度,使 PTH-Ala 和 PTH-His 保留时间间隔为 0.5～1.0min,便于 PTH-PEC 的分离。

（三）结　　果

① 参见图 1.7，还原型 SOD（左图）与天然 SOD（右图）的肽谱，上图是 220nm 检测，下图是 254nm 检测，仅显示含巯基的肽。

② 含巯基肽的序列分析

按 HPLC 图谱洗脱次序依次命名肽 1、肽 2、肽 3。

肽 1：AVCVLK（4～9）；

肽 2：LACGVIGIAK（53～62）；

肽 3：GDTVVVTGSITGLTEGDHGFHVHQFGDNTQGCTSAGPHFNPLSK（113～156）；

由此得出：肽 1 中 Cys-6 是游离的一巯基，而肽 2 中的 Cys-55 和肽 3 中的 Cys-144 形成二硫键。

（四）提　　示

①修饰天然 SOD 所需的乙烯吡啶量大大多于修饰还原型 SOD 所需的量，可能原因是 SOD 在还原过程中蛋白质结构趋于松散，半胱氨酸相对稳定，易于乙烯吡啶结合。

②调节序列仪中 120A HPLC 分离 PTH 氨基酸洗脱液 A（5% 四氢呋喃）中 HAc-NaAc 的离子强度，使 PTH-Ala 和 PTH-His 保留时间间隔为 0.5～1.0min，便于 PTH-PEC 的分离。

二、含芳香族氨基酸肽的分离纯化和鉴定

（一）引　　言

蛋白质在 280nm 有特征性的高吸收，这主要是芳香族氨基酸（色氨酸、酪氨酸和苯丙氨酸）的贡献。蛋白质在酶解后的肽谱中，检测波长不用 280nm，而采用 214～220nm，这是由于肽键的强吸收缘故。肽谱中不含芳香族氨基酸的肽，在 280nm 下，基本上没有吸收峰，但含有色氨酸或酪氨酸的肽，则在 280nm 下却表现出强的吸收峰。据此，我们可以辨别含芳香族氨基酸的肽。

此外，含芳香族氨基酸肽的二阶微分光谱也和一般肽的二阶微分光谱有明显的差别。有报道在碱性条件下（pH 12.5），Tyr 离子化的二阶微分光谱与 Trp 的二阶微分光谱完全不同，特别在 260nm[36]。但这样的条件对酸性的 HPLC 反相柱显然是不合适的。我们用不含 Trp 和 Tyr 的 19 肽（从血红蛋白酶解物中分离纯化），含有 Trp 的 5 肽胃泌素，含有 Tyr 的合成 7 肽和同时含有 Trp 与 Tyr 的 LH－RH 作为模型肽，进行了 HPLC 分析和在线二阶微分光谱分析，结果见图 2.4[37]。

由图谱可见，不含 Trp 和 Tyr 残基的肽，二阶微分光谱为一无序图谱，没有 Trp 和 Tyr 的特征峰形（图 2.4a）；含 Trp 的肽的二阶微分光谱在 290nm 处有特征峰（图 2.4b）；含 Tyr 肽的二阶微分光谱在 283nm 处有特征峰（图 2.4c）；同时含有 Trp 和 Tyr 的肽的二阶微分光谱在 290nm 和 283nm 处都有特征峰（图 2.4，d）。据此也可以辨别含不同芳香族氨基酸的肽[36]。

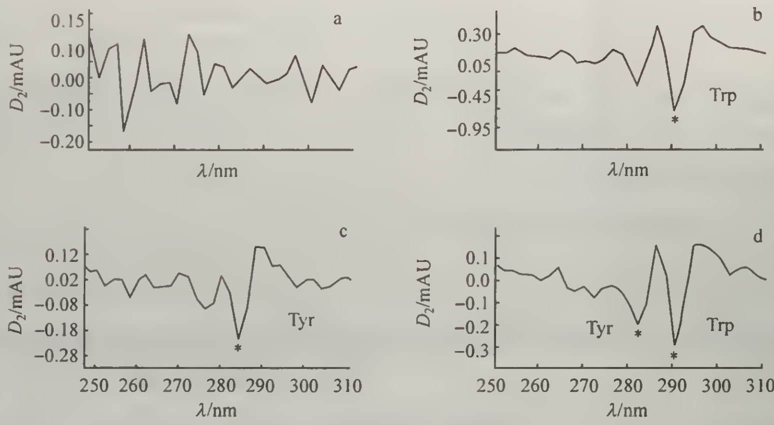


图 2.4　模型肽的在线二阶微分光谱分析

a.19 肽;b.5 肽胃泌素;c. 合成 7 肽;d.LH-RH。

(二)不同波长测定检出含芳香族氨基酸的肽

由于芳香族氨基酸在 280nm 附近有强吸收,但它们的强度差别也很大,一般是 Trp≫Tyr≫Phe。基于此,我们可以比较同一肽图在 214nm 和 280nm 下的吸收,可以方便地判断哪一个肽可能含有芳香族氨基酸。

1. 材料

新型白细胞介素-2(Ser-125):上海医药工业研究院中试组提供。

2. 方法

① 酶解:TPCK-胰蛋白酶为本实验室自制。IL-2:TPCK-Trypsin=50:1(*W/W*),碳酸氢铵缓冲液,pH 8.0,37℃酶解 4h,加 TFA 至 pH 2～3 停止反应。

② HPLC:Beckmam System Gold 系统,250mm×4.6mm I.D. RP-18 柱。梯度洗脱条件为 0.1%TFA/H_2O-0.085%TFA/CAN,60min。

③ 蛋白质/多肽序列分析:ABI 477A 序列仪,常规分析。

3. 结果

参见图 1.9,这是白细胞介素-2 的酶解肽谱,上图是 214nm 检测,下图是 280nm 检测。由此可方便地判定下图中的高吸收峰含有芳香族氨基酸[38]。

(三)二阶微分光谱检出含芳香族氨基酸的肽

1. 材料

19 肽:由血红蛋白 β 链酶解后从 HPLC 制备,No. 41～59。

Phe-Phe-Glu-Ser-Phe-Gly-Asp-Leu-Ser-Thr-Pro-Asp-Ala-Val-Met-Gly-Asn-Pro-Lys

5 肽胃泌素:上海丽珠东风生物技术有限公司产品。

Boc-β-Ala-Trp-Thr-Pro-Phe-NH_2

合成 7 肽:崔大敷教授合成并赠送。

Phe-Phe-Tyr-Thr-Pro-Arg-Thr

LH-RH:p-Glu-His-Trp-Ser-Tyr-Gly-Leu-Arg-Pro-Gly-NH_2

重组人表皮生长因子(hEGF):甘人宝教授赠送。

新型人白细胞介素-2(Ser-15):天元生物技术有限公司提供。

2. 方法和仪器

①HPLC:HP 公司 HPLC 1090M Series II,具有 DAD 检测器和相应化学工作站软件,可在线对图谱各峰进行二阶微分光谱分析。HP-C18 柱(100mm×2.1mm I. D.);TEA-ACN 梯度(均为 HPLC 级产品);mAU 为毫吸收单位。

②酶解条件:TPCK-胰蛋白酶为本实验室自制。HEGF:胰蛋白酶=50:1(*W/W*),0.1mol/L 碳酸氢铵缓冲液,pH 8.3,37℃酶解 18h 后,用稀 TFA 调节至 pH 2.0 左右上柱分析。

③氨基酸分析:适量肽于 110℃下,恒沸点 HCl 水解 22h(Waters 水解工作站)后,于 Beckman 6300 氨基酸组成分析仪上分析。

④蛋白质序列分析:ABI 477A/120A 蛋白质序列仪上,按常规方法进行。

3. 结果

(1) hEGF 的肽谱及含 Trp 肽的分析

图 2.5a 图是 hEGF 的酶解产物的肽谱;对肽谱中每个峰进行二阶微分光谱分析,只有标星号的肽,在 290nm 处有一明显的负峰,是 Trp 肽的特征二阶微分光谱。见图 2.5b。

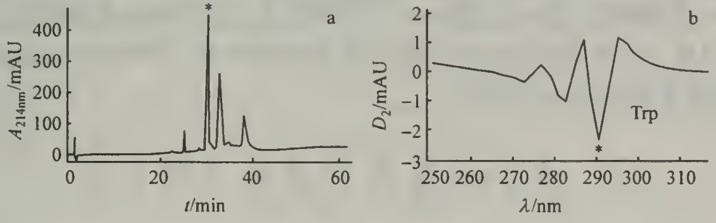


图 2.5　hEGF 的胰蛋白酶酶解肽谱(a)及其峰 2 的二阶微分分析(b)

根据 EGF 的一级结构,它仅在 C 端肽中有两个 Trp 残基。收集星号肽进行氨基酸分析,仅 Glu、Leu、Arg,其比值近于 1,Trp 在酸水解时被破坏。对肽的序列分析结果为:-Trp-Trp-Glu-Leu-Arg。由此判定此肽是 EGF 的 C 端肽。

(2) 新型 IL-2 的肽谱及其相关分析

新型 IL-2 是重组蛋白,其突变点是原来的 Cys-125 改变成 Ser-125。将新型 IL-2 用 TPCK-胰蛋白酶酶解后,进行了 RP-HPLC 分离分析。因为 IL-2 中只有一个 Trp 残基位于 121 位。

肽谱的二阶微分光谱扫描判定仅星号肽具有 Trp 肽的特征光谱,即此肽是含有 Trp 的 C 端肽(No. 121~133)(图 2.6)。

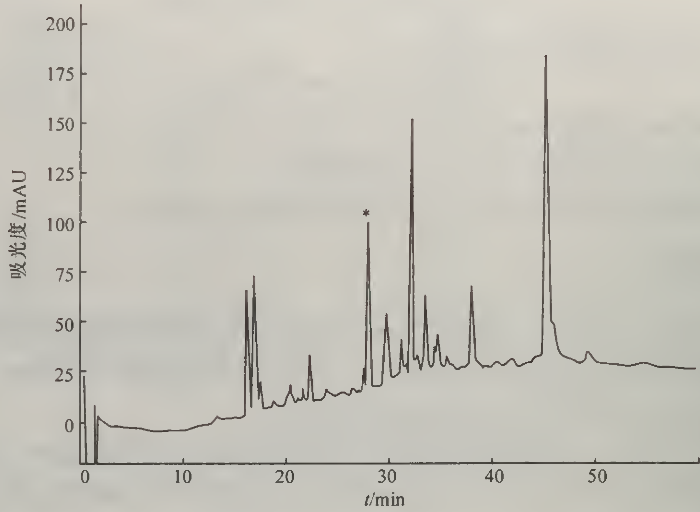


图 2.6　IL-2 的酶解肽谱及其二阶微分光谱分析

收集此肽进行序列测定，结果是：

121　　　　　　125　　　　　　　130　　　　133
Trp-Ile-Thr-Phe-Ser-Glu-Ser-Ile-Ile-Ser-Thr-Leu-Thr

更证实了二阶微分光谱结果的可靠性。

(3) 讨论

二阶微分光谱分析具有简便、快速的优点，不必增添任何试剂处理并能在线检测。但此法也有一定的局限性，Trp 和 Tyr 的二阶微分光谱都有次峰谷，Trp 除在 290nm 有特征峰外，在 282nm 有次峰谷；Tyr 在 282nm 有特征峰外，在 276nm 有次峰谷。它们且有重叠现象，所以判别 Trp 肽中含有 Tyr 的存在是件困难的事，特别当 Trp∶Tyr＝1∶1 时，因为 Trp 的光吸收系数比 Tyr 大得多。

三、含 Lys-肽的分离纯化和鉴定

赖氨酸的 ε-氨基可以和磷酸吡哆醛(PLP)形成 Schiff 碱，该衍生物在 325nm 有特殊的吸收，利用这一性质可找出含有 Lys 残基的肽段[39]。

(一) 引　言

蛋白质中的赖氨酸残基常常与生物活性密切有关。例如醛缩酶(ALD)中的 Lys-227 和同位素标记的磷酸二羟丙酮反应后，经 $NaBH_4$ 还原可以形成 Schiff 碱。此时酶分子上可测定到等克分子数的放射性，同时酶的催化活性丧失。

用同位素标记肽比较麻烦，Shapiro[40]则发现醛缩酶中 Lys-117 和磷酸吡哆醛(PLP)反应，并经 $NaBH_4$ 还原也可以形成 Schiff 碱，酶活性受抑制，但修饰后的醛缩酶仍可和底物结合。由此推论 Lys-107 是一次要活性部位。

用 PLP 作用的优点是避免用同位素，PLP 修饰在 Lys 残基上形成的肽，在 325nm 有

一特殊吸收[40]。因此，这种技术对制备和分离 Lys-肽具有普遍意义。

（二）材料与试剂

醛缩酶：本实验室制备。

PLP：Sigma 公司产品。

$NaBH_4$：Aldrich 公司产品。

ACN：Fluka 公司产品。

TFA：Merck 公司产品。

（三）方 法

①PLP 修饰：2mg ALD（12.5nmol）和 8μl 10mmol/L PLP 在 2ml 10mmol/L，pH7.5 Tris-HCl 缓冲液中室温避光反应 30min，pH 调至 6.5 后加入 20μl 0.5mol/L $NaBH_4$（10μmol），0℃反应 10min。对重蒸水透析。在分光光度计上全波扫描来计算 PLP 的结合量，以 $\varepsilon_{280nm}=1.46\times10^5$L/mol 定量 ALD 蛋白含量，以 $\varepsilon_{325nm}=1.015\times10^4$L/mol 定量 PLP 衍生物的含量。

②TPCK-胰蛋白酶水解：1mg/ml ALD 在 100mmol/L pH8.0 NH_4HCO_3 中，37℃反应过夜。底物与酶重量比为 50∶1。

③酶解产物的分离：酶解产物经 ABI 150A HPLC 分离系统，柱为 Spheri-5（C18，50mm×1.0mm I.D.）。流动相 A 为 0.1%TFA；流动相 B 为 0.085%TFA－70%CAN。洗脱条件：0%～25%B 20min；25%～75%B 60min；75%～100%B 10min。流速为 0.1ml/min。

④肽纯度鉴定：在 ABI 270A 型毛细管电泳上分析。毛细管长 75cm（至检测器距离 50cm），内径 50μm。电泳条件为 20mmol/L pH2.5 柠檬酸缓冲液，电压 20kV，电流 15μA。

⑤氨基酸序列分析：在 ABI 477A/120A 蛋白质/多肽序列仪上常规进行。但在 120A PTH 分析系统中改变了乙酸盐的浓度，使 PTH-Ala 与 PTH-His 洗脱时间相差约 1.5min，这样可将 PTH-Ala 与 PTH-Lys*（经 PLP 修饰）能分离开。

（四）结 果

（1）PLP 结合位点肽段的分析

图 2.7 为 ALD 经 PLP 修饰后的胰蛋白酶酶解图，其中图 2.7a 在 220nm 波长检测，图 2.7b 在 325nm 波长检测。

分别收集图 2.7 b 中的峰 1，2，3，经毛细管电泳鉴定纯度，如纯度不够，则在 HPLC 上稍加改变洗脱梯度再纯化。

（2）含 Lys-肽的序列分析结果

101 107 110

肽 1：G-G-V-V-G-I-K*-V-D-K

28 41 42

肽 2：G-I-L-A-A-D-E-S-T-G-S-I-A-K*-R

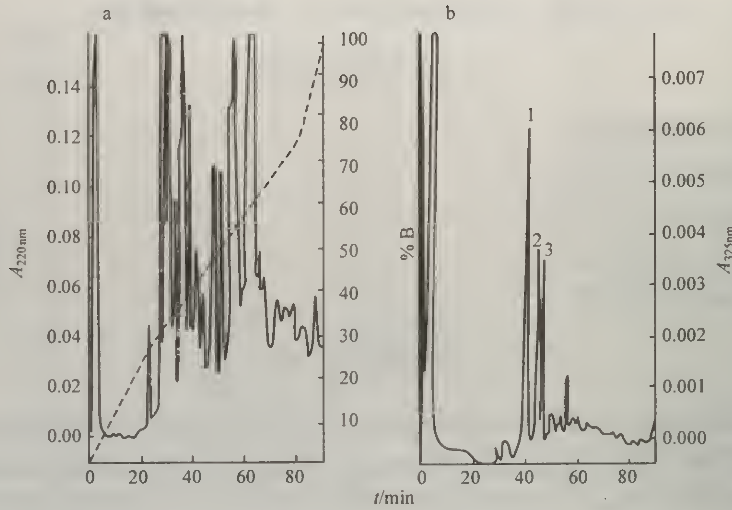


图 2.7 PLP-修饰的醛缩酶经胰蛋白酶水解产物的 HPLC 图谱

87 91 98

肽 3:A-D-D-G-R* -P-F-P-Q-V-I-K

(夏其昌)

参 考 文 献

1 Stone KL *et al*. RP-HPLC separation of sub-nanomole amounts of peptides obtained from enzymatic digest. In: Mant C, Hodges RS eds. HPLC of peptides and proteins: separation, analysis and conformation. Boca Raton: Florda, CRC Press, 1991. 669~677

2 Matsudaira P. Sequence from picomole quantities of peptides electroblotted onto PVDF membranes. J Biol Chem, 1987, 262:10035

3 Yuen SW, Chui AH, Wilson KJ *et al*. Microanalysis of SDS-PAGE electroblotted proteins. Biotechniques. 1989, 7(1):74~83

4 夏其昌,超滤法及其应用.见:鲁子贤主编.蛋白质和酶学研究方法,北京:科学出版社,1989. 42~52

5 Fujita VS, Black SD, Tarr GE *et al*. On the amino acid sequence of cytochrome P-450 isozyme 4 from rabbit liver microsomes. Proc Natl Acad Sci U S A. 1984, 81(14):4260~4264

6 Gross E, Witkop B. The hetrogeneity of the S peptide of bovine pancreatic ribonuclease A. Biochemistry, 1967, 6:745~748

7 Xia QC, Chang D, Blacher R *et al*. The primary structure of the COOH-terminal half of cholera toxin subunit A1 containing the ADP-ribosylation site. Arch Biochem Biophys, 1984, 234(2):363~370

8 Omenn GS, Fontana A, Anfinsen CB. Modification of the single tryptophan residue of staphylococcal nuclease by a new mild oxidizing agent. J Biol Chem, 1970, 245:1895~1902

9 Fontana A. Modification of tryptophan with BNPS-Skatole (2-(2-nitrophenylsulfenyl)-3-methyl-bromoindolenine). Methods Enzymol, 1972, 25:419~423

10 Mahoney WC. Hermodson MA. High-yield cleavage of tryptophenyl peptide bonds by o-iodosobenzoic acid. Biochemistry, 1979, 18:3810~3814

11 Landon M. Cleavage at Asparyl-Prolyl bonds. Methods Enzymol, 1977, 47:145~149

12 Bornstein P, Balian G. Cleavage at Asn-Gly bonds with hydroxylamine. Methods Enzymol, 1977, 47:132~145

13 Jekel PA, Weijer WJ, Beintema JJ. Use of endoproteinase Lys-C from Lysobacter enzymogenes in protein sequence analysis. Anal Biochem, 1983, 134(2):347～354

14 Masaki T, Tanabe M, Nakamura K *et al*. Studies on a new proteolytic enzyme from A chromobacter lyticus M497-1. I. Purification and some enzymatic properties. Biochim Biophys Acta, 1981, 660(1):44～50

15 Levy M, Fisman L, Shenkein I. Mouse submaxillary gland protease. Methods Enzymol, 1972, 19:672

16 Drapeau GR. Substrate specificity of a proteolytic enzyme isolated from a mutant of *pseudomonas fragi*. J Biol Chem, 1980, 255:839

17 Ishi S *et al*. An asparaginyl endopeptidase purified from Jack bean seeds. J Protein Chem, 1990, 9:294

18 Houmand J, Drapeau GR. Staphylococcal protease: a proteolytic enzyme specific for glutamoyl bonds. Proc Natl Acad Sci USA, 1972, 69:3506

19 Brown WE, Wold F. Alkyl isocyanates as active-site specific reagent for serine proteases, reaction properties. Biochemistry, 1973, 12:828

20 Gross E. The cyanogen bromide reaction. Methods Enzymol, 1967, 11:238

21 Fontana A. Modification of tryptophan with BNPS-skatole. Methods Enzymol, 1972, 25:419

22 Mahoney WC, Hermodson MA. High yield cleavage of tryptophanyl peptide bonds by o-iodosobenzoic acid. Biochemistry, 1979, 18:3810

23 Fontana A, Dalzoppo D, Grandi C *et al*. Cleavage at tryptophan with o-iodosobenzoic acid. Methods Enzymol, 1983, 91: 311－318

24 Schecher I *et al*. Selective chemical cleavage of tryptophenyl peptide bonds by oxidative chlorination with N-chlorosuccimimide. Biochemistry, 1976, 15:5071

25 Mitchell WM. Cleavage at Arginine residues by clostripain. Methods Enzymol, 1977, 47:165～170

26 Lundblad RL, Kingdon HS, Mann KG. Related articles thrombin. Methods Enzymol, 1976, 45:156～176

27 Carter P. Site-specific proteolysis of fusion proteins. In: Ladisch MR. *et al* eds. Protein Purification. American Chemical Socity, 1990, 181～193

28 Richard FM. Vithayathil PJ. The preparation of subtilisin-modified ribonuclease and the separation of the peptide and protein component. J Biol Chem, 1959, 234:1459～1465

29 Wellner D, Panneerselvam C, Horecker B L. Sequencing of peptides and proteins with blocked N-terminal amino acids N-acetylserine or N-acetylthreonine. Proc Natl Acad Sci USA, 1990, 87:1947

30 Mitta M, Asada K, Uchimura Y *et al*. The primary structure of porcine liver acylamino acid-releasing enzyme deduced from cDNA sequences. J Biochem, 1989, 106(4):548～551

31 Allen G. Sequencing of proteins and peptides. 2nd ed. Elsevier, 1989, 295～326

32 Aebersold R. Mass spectrometry of proteins and peptides in biotechnology. Curr Opin Biotechnol, 1993, 4:412～419

33 屠红旻,夏其昌.毛细管电泳测定牛红细胞超氧化物歧化酶的巯基位点.生物化学杂志,1992,8:570～573

34 Steinman HM, Naik VR, Abernethy JL. Bovine erythrocyte superoxide dismutase. J Biol Chem, 1974, 249:7326～7338

35 Abernethy JL, Steinman HM, Hill RL. Bovine erythrocyte superoxide dismutase. J Biol Chem, 1974, 249:7339～7344

36 Balestrieri C, Colonna G, Giovane A *et al*. Second-derivative spectroscopy of proteins: studies on tyrosyl residues. Anal Biochem, 1980, 106(1):49～54

37 张丽华,屠红旻,邵晓霞等. HPLC在线二阶微分光谱法检测基因工程产品的C端肽.科学通报,1995,40(14): 1314～1317

38 吴高德等.新型人白介素-2突变点的分析鉴定.生物工程学报,1992,8:314～317

39 屠红旻,夏其昌.蛇肌与兔肌醛缩酶磷酸吡哆醛结合位点及作用机制的比较研究.生物化学与生物物理学报,1993,25:175～180

40 Lai CY, Nakai N, Chang D. Amino acid sequence of rabbit muscle aldolase and the structure of the active center. Science, 1974, 183(130):1204～1206.

附录一　微量蛋白质实验室注意事项

(1) 环境

由于是微克级的操作,实验室内的温度、灰尘、纤维、皮肤和头皮屑都会影响结果。

(2) 吸附

玻璃容器的表面吸附,由于量少,会使大部分物质被吸在壁上. 在溶液中的量趋于零。

用透析袋,少量蛋白质(100μg)会因吸附作用而导致严重损失。建议用微量透析仪器,或用小的凝胶过滤柱脱盐,甚至用挥发性溶剂,避免脱盐。

推荐储存少量肽(<10nmol)在小的玻璃管(5mm×50mm)中,避免转移和表面吸附引起的损失。

冻干样品时,如果用大的器皿,则用硅烷化处理;有时候,宁可存放样品在聚丙烯管中。

(3) 样品加样和转移

小量样品采用称量是不精确的,因为有的样品含盐易吸湿。蛋白质/肽溶在适合的小量溶剂中(例 50μl),用 5~10μl 进行分析,而且建议用玻璃的微量注射针筒,这比用微量可调加样器精确得多。

吡啶-乙酸-水是相当好的蛋白质/肽的溶剂,但吡啶在 280nm 有很高的光吸收,影响到紫外定量和 HPLC 的检出。此时,较大的酸性肽可溶在 NH_4OH,NH_4HCO_3,有些蛋白质/肽在此溶剂中不溶解,则可用酸(70%甲酸,或 TFA),并尽快在真空干燥后于 −20℃ 保存。为防止在过程中空气泡的喷出而使样品丢失,开始可用水泵抽 10min,直至溶解在水中的空气泡逸出后,再用油泵真空干燥。

(4) 反应

小体积的溶液,先用微量离心机使样品集中在底部,加入试剂后迅速密闭防止小体积溶剂的挥发,然后涡旋,样品和转移均用微量注射器。

附录二　常用试剂处理

(1) 水

在我们实验室采用 Mini-Q 水。必须指出:即使用 Mini Q 水也要新鲜,长期存放的 Mini Q 水,在 HPLC 空白时,也会有峰出现。

(2) 溴化氢(cyanogen bromide,BrCN)

应新鲜配制,完全无色才能用。平时密闭保存在 −20℃。

(3) 盐酸胍(Guanidium chloride,Gu. HCl)

推荐用 Schwarz/Mann Biotech 公司的超纯级产品。纯化方法见 Nozaki,Y.,Methods Enzymol.,1972;26:43~50

(4) 6mol/L 盐酸

浓盐酸加等体积水,重蒸三次,去除前后馏分。此产品通常认为是 6mol/L

(或 5.7mol/L)的 HCl。

(5) 尿素

尿素溶液在放置时，特别在碱性 pH 下，会有氰化胺(ammonium cyanate)积累，故应用新鲜配制尿素溶液。

氰化胺除去的方法：溶液通过一混合离子树脂的小柱，1L 8mol/L 尿素溶液需 10g 树脂，开始流出柱的两个柱体积的尿素丢弃不用。

第三章　氨基酸组成分析

氨基酸是一种小分子的两性化合物，分子质量在 75～200Da 之间，其化学通式为 $\mathrm{R{-}\underset{\underset{\displaystyle NH_2}{|}}{CH}{-}COOH}$ ，在生物体内出现的氨基酸都是 *L* 型，仅在少数微生物来源的多肽中出现 D 型氨基酸。常见的 20 种氨基酸结构式见表 1.1。

氨基酸分析技术的发展始于 1820 年，由 Braconnot 最早将甘氨酸从白明胶水解物中分离出来，但直到 21 世纪初，用化学方法或微生物方法测定氨基酸，仍是一项持续数周或数月的繁琐的工作。

色谱技术的引入为氨基酸分析打开了一扇新的大门。1941 年，Martin 等运用色层法分析氨基酸，主要采用乙酰化氨基酸在硅胶柱上或滤纸上半定量；随后，Moore、Stein 应用离子交换柱直接分离游离的氨基酸，在 1958 年，Spackman、Stein、Moore[1] 制造了自动化的氨基酸分析仪，使氨基酸定量分析进入了一个崭新的阶段，至今仍是经典的方法。

第一节　氨基酸组成分析的目的

氨基酸是维持生物体正常机能的必需物质，氨基酸分析在生命科学、制药、食品、饲料和临床检验等领域具有非常重要的意义。现代分离提纯技术的发展使蛋白质操作微量化，但也给定量带来了困难，一般很难通过常规的称量或测蛋白质溶液在 280nm 的光吸收值来准确定量蛋白质。所以，如果在蛋白质酶解或测序前，取蛋白质样品的一部分进行氨基酸组成分析，根据结果便可以推算出蛋白量的可靠值。另外，在蛋白质测序中有时遇到测不出结果的情况，一种可能是蛋白质的 N 端封闭，另一种可能则是样品本身不是蛋白质或绝大部分是非蛋白质物质，解决这个问题的很好途径便是做氨基酸组成分析以确定样品的成分。

除了蛋白质研究和重组蛋白需要测定氨基酸组成外，医学上也需要测定血液或各种体液中的游离氨基酸。

第二节　蛋白质的水解方法

氨基酸分析的第一步是将蛋白质或多肽水解，常规方法是采用 5.7mol/L 盐酸在真空状况下于 110℃ 水解 24h，也有在 150℃ 下快速水解 90min[2]。Woodword[3] 等人提议用微波辐射法（20～25min，150℃）来水解蛋白质和多肽。如果是测定生理样品中的游离氨基酸，在样品预处理时主要是通过强酸、有机溶剂、盐析、加热、超滤和预柱等方法将大分子蛋白质除去，评价除去蛋白质方法的有效性应从去除蛋白质效果和待分析样品的回收两方面来评价

我们实验室目前采用的是气相水解法，操作简便，样品的损失小，回收高。Stein[4]在1987年提出气相水解法，而后Waters公司有商品化仪器（水解工作台）出售（图3.1a）。将100～500pmol的蛋白质或肽转移到水解指管中（图3.1b）冷冻抽干，在水解反应器中（图3.1b）加入5.7mol/L盐酸200μl，含2%苯酚（m/V），将指管置于反应器中，抽真空并充氮气反复三次（图3.1c，真空控制也是实验重复性的一项重要因素）110℃保温22h。其中，苯酚是作为一种抗氧化剂来防止半胱氨酸（Cys）和甲硫氨酸（Met）的氧化。如果遇到一些不易裂解的肽键，如Val—Val，Ile—Leu等，水解时间可适当延长至72h，但长时间水解会部分破坏含羟基的氨基酸[丝氨酸（Ser），苏氨酸（Thr），酪氨酸（Tyr）]。



图3.1　氨基酸水解工作台

a.水解台：高温水解室（1）；温度指示（2）；真空表（3）；真空操作部位（4）；开关（5）；冷颈（6）。b.水解指管（7）和水解反应器（8）。c.真空操作部位：通氮（9）；真空（10）。d.水解室可容纳4个水解反应器

影响氨基酸分析结果的三大因素是：水解、衍生、色谱。水解是第一步，水解的完全与否对结果影响很大。在我们实验中采用一些标准蛋白质，如细胞色素c（Cyt c）、核糖核酸

酶(RNase)或胰岛素(insulin)为标样进行水解、衍生和色谱分析来计算回收,从而考验蛋白质分析的可靠程度和评价方法的适用性。

表 3.1　一些标准蛋白质的氨基酸组成分析

氨基酸	胰岛素		马心细胞色素 c	
	理论值	实验值	理论值	实验值
Asp	3	2.8	8	8.3
Thr	1	1.0	10	7.9
Ser	3	2.8		
Glu	7	7.0	12	13.5
Pro	1	1.2	4	3.8
Gly	4	4.1	12	10.7
Ala	3	3.0	6	5.9
Cys*	6	2.3		
Val	5	4.0	3	4.0
Met			2	2.4
Ile	1	0.7	6	6.2
Leu	6	5.8	6	6.2
Tyr	4	3.8	4	3.3
Phe	3	3.0	4	3.6
His	2	2.0	3	3.4
Lys	1	1.2	19	20.1
Arg	1	1.3	2	2.4

* 未作任何处理。

但由于各种蛋白质含有的氨基酸多少不同,并且每一种氨基酸的量也不一样,有的氨基酸含量可多到十几个,有的氨基酸含量只有 1～2 个,这影响到氨基酸的回收。生物应用公司曾设计过一种“测试肽”,即人工合成 18 种不同氨基酸的肽(除 Gln、Asn 外),这样每种氨基酸出现的频率都是 1,用此肽来校正回收,看来机会均等,是一种可以采纳的方法。我们也对合成此肽作了尝试,但在实际样品中,不太可能出现这种理想情况。

测试肽的结构:

H-Cys-Pro-Asp-Phe-Gly-His-Ile-Ala-Met-Glu-Leu-Ser-Val-Arg-The-Trp-Lys-Tyr-OH

第三节　特殊氨基酸的保护和定量

不同条件下,各种氨基酸的回收有所不同,但色氨酸均遭破坏,通常解决方法有以下 5 种。

1. 水解酸中加添加剂

这些添加剂包括巯基乙酸[4]和 β-巯基乙醇[5],我们通过实验认为液相水解时加入巯基乙醇对色氨酸保护效果较好,色氨酸的回收可达 80%。有报道 Perkin-Elmer 公司针对其产品 ABI 420A 的特点,推出试剂十二烷基硫醇[6],专门保护色氨酸,据称回收率达 97%,但由于其试剂是滴在加样孔的砂芯玻璃上使其在水解时造成一种还原环境来保护色氨酸的,只适用在该公司特定的仪器上,不能在其他仪器上推广。

2. 有机酸

3mol/L 巯基乙磺酸(mercaptoethanesulfonic acid)[7]或 4mol/L 甲磺酸(methane-sulfonic acid)[8]在水解时对色氨酸有一定的保护作用。

3. 碱

用氢氧化钠[9]和氢氧化钡[10]代替酸水解,可保护色氨酸不被破坏。

4. 酶

蛋白酶水解法有时也有应用,其优点是条件温和,对天冬酰胺和谷氨酰胺及色氨酸均无破坏作用,但酶水解法往往不完全,必须用酸水解法校正。

5. 光谱法

分光光度法[11]和二阶微分光谱法[12]可测定色氨酸,前者应用较早,在蛋白质分子中,如不含胱氨酸和其他有干扰的发色基团的氨基酸,根据其在 6mol/L 盐酸胍的中性溶液中的紫外吸收,测定色氨酸和酪氨酸的含量,此经验公式只适用于色氨酸对酪氨酸的比值为 0.2～1 时的情况。后者常被用来确定蛋白质中芳香族氨基酸的含量,计算公式复杂。Hewlett Packard 公司推出的 DAD 检测器配合其化学工作站软件,能较方便地在线检测并定量肽段中的色氨酸。

虽然蛋白质和多肽的氨基酸分析法对大多数氨基酸是快速、准确和灵敏的,但在半胱氨酸定量时仍遇到很大困难,解决这个问题的途径有两种,一种是还原烷化法,另一种是氧化法。

(1) 还原烷化法[8]

产生能抗水解的半胱氨酸衍生物。烷化试剂包括 4-乙烯吡啶和碘代乙酸。但这些方法有缺点,为了确保试剂接近巯基,还原和氧化通常在变性剂存在下进行,多余试剂和变性剂要用 HPLC、沉淀法或透析去除,每一步都可能造成样品的损失,特别是对于小肽和微量样品,而且烷化反应如果不仔细控制条件,可能导致半胱氨酸以外其他残基的衍生。

(2) 氧化法

氧化反应被用来转化半胱氨酸和胱氨酸成为半胱磺酸,过甲酸[13]氧化反应的缺点是它会影响其他氨基酸特别是甲硫氨酸会转变成甲硫氨酸砜,酪氨酸会降解。为克服这一缺点,必须准备能分析两次的样品量,一次提供半胱氨酸数据,另一次则提供其他氨基酸的组成数据。

Barkholt 和 Jensen[14]提出使用二巯基化合物(如 DTDPA、DTDBA)作为酸水解步骤中的保护剂,他们发现半胱氨酸和这些化合物形成稳定的衍生物,是一种双巯基混合物,衍生的半胱氨酸残基能通过离子交换层析与其他氨基酸分离,并用柱后衍生法检测到。二巯基试剂对其他氨基酸无影响,样品只须稍加处理,通常取溶解于 0.2mol/L 氢氧化钠中的 1% DTDPA 20μl,与 1～5nmol 蛋白质反应,再用 20μl 异丙醇淋洗管壁,抽干后水解。

但以上两种试剂在分析 PTC-AA 时常会干扰层析谱图。Hoogerheide[15]等通过改变二巯基二酸的链长度而改变衍生物在层析谱图中的位置,发现了另一种半胱氨酸保护

剂——DTDGA,同时他们认为,无论是气相水解还是液相水解,二巯基二酸必须直接接触样品,否则衍生物产率下降,导致半胱氨酸定量不准。

第四节　氨基酸组成原位分析

如果蛋白质样品采用电泳法(如 SDS-PAGE)分离,则很难从凝胶中洗脱下来,可采用电印迹法把样品转移到 PVDF 膜上,然后在 PVDF 膜上直接进行盐酸水解和氨基酸分析,称之为原位(*in situ*)分析。因为通常所用聚丙烯酰胺凝胶分离系统中含大量甘氨酸和 Tris 盐,易干扰氨基酸原位分析,故改用 CAPS 缓冲系统以减少影响。另外,在膜上做氨基酸组成分析时可以做一个空白对照,具体操作时将含蛋白条带的 PVDF 膜割下,另取一条相同面积的转移膜作为空白,分别置于水解指管中,加入 70μl 含 7%(m/V)巯基乙酸的 6mol/L 盐酸以防止抽真空时膜片的漂出,将水解指管放入水解反应器中,在反应器底部加入 200μl 含 7% 巯基乙酸的 6mol/L 盐酸,110℃ 真空水解 24h,水解结束后用 200μl 含 30% 甲醇的 0.1mol/L 盐酸反复三次将氨基酸从 PVDF 膜中抽提出来,以便进行下一步的分析。

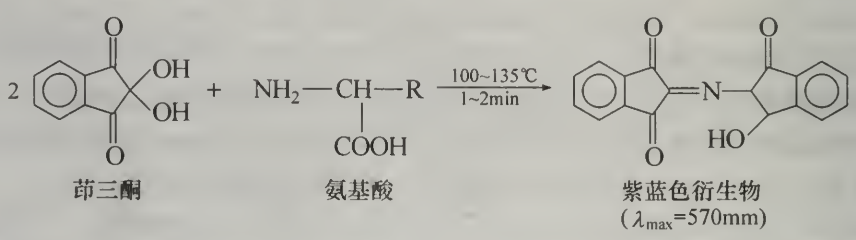
第五节　衍生方法及原理

从衍生角度看,氨基酸分析可以分为柱后反应法和柱前衍生法两大类。前一种方法是将游离氨基酸经过色谱柱分离后,各种氨基酸再与显色剂(茚三酮、荧光胺、邻苯二甲醛)作用,这种方法比较稳定,对样品预处理要求低,容易定量和自动化操作,不足之处在于检测灵敏度不高(100pmol 以上),分析时间长(蛋白质水解液需 1h 而某些生理样品则需 4h 以上)。另一种方法是将氨基酸和化学耦联试剂先反应生成氨基酸的衍生物,然后再用色谱柱将各种衍生物分离,直接检测衍生物的光吸收或荧光发射,此法可检测 OPA-、PTC-、PTH-、DABS-、Dansyl-和 DABTH-氨基酸,分析灵敏度高,可利用 HPLC 进行氨基酸分析,缺点是有的衍生物不稳定,衍生试剂可能干扰氨基酸检测。

以下是几种常见的氨基酸分析方法。

(1) 茚三酮法

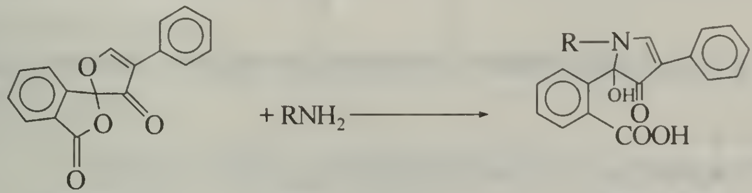
McCaldin 曾详细研究了茚三酮与氨基酸反应的机理,Spackman 等人最早将茚三酮试剂运用到氨基酸组成分析上。离子交换层析后,各种氨基酸与茚三酮反应形成紫色化合物,最大光吸收在 570nm,摩尔消光系数 2×10^4,茚三酮和二级胺(脯氨酸和羟脯氨酸)形成黄色产物,在 440nm 检出,摩尔消光系数 3×10^3。



目前此法的灵敏度可达 100pmol，但茚三酮试剂易氧化，必须隔绝空气避光保存，试剂本身黏性大，需要有个柱后混合器才能与氨基酸充分反应，对仪器要求较高。

(2) 柱后荧光胺法

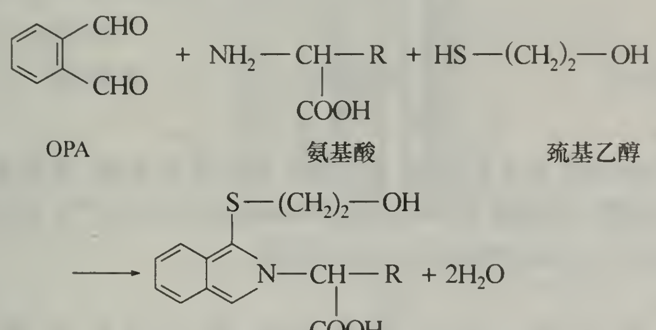
20 世纪 70 年代合成了荧光胺，它能在室温下迅速和一级胺发生反应，其荧光产物的激发波长 Ex390nm，发射波长 Em475nm。次级胺可以用 N-Chlorosuccinimide (NCS)氧化生成相应的初级胺后再与荧光胺反应而检出[16]。



荧光胺　　　　一级胺　　　　荧光产物

(3) 邻苯二甲醛(OPA)法

OPA 最早是被作为柱后衍生试剂而引进的，后来也逐渐应用于柱前反应中。OPA 能在还原剂巯基乙醇存在下和一级胺产生具有很强荧光的异吲哚衍生物，反应迅速，在室温下 1min 即可完成，反应络合物的 Ex340～360nm，Em455nm，灵敏度比茚三酮法高，在 OPA 中加入 SDS 或表面活性剂 Brij-35，不但能提高赖氨酸和羟赖氨酸的稳定性，而且能提高它们的荧光度。缺点是不能检测次级氨基酸，但可用稀的 Na-hypochlorite[17]在碱性条件下氧化成相应的一级胺，然后再与 OPA 反应。

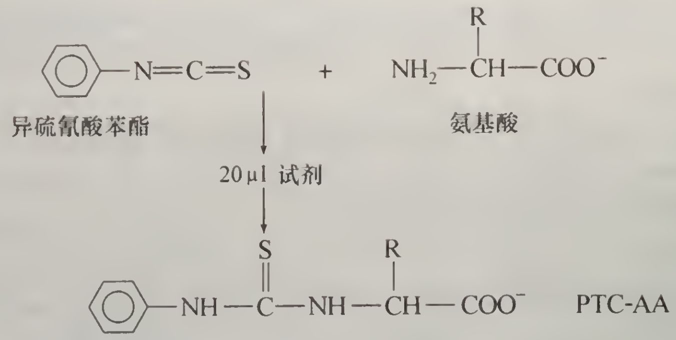


异吲哚衍生物

OPA/FMOC 作为柱前衍生剂能同时与一级胺和二级胺反应，室温下只需几分钟，Hewlett Packard 公司改用巯基丙酸(3-MPA)[18]代替巯基乙醇。由于 3-MPA 的加入，异吲哚衍生物疏水性降低，使出峰时间提前，过量的 FMOC 及其分解产物的出峰时间在最后，被洗脱出的次级氨基酸之后，不干扰分析结果。

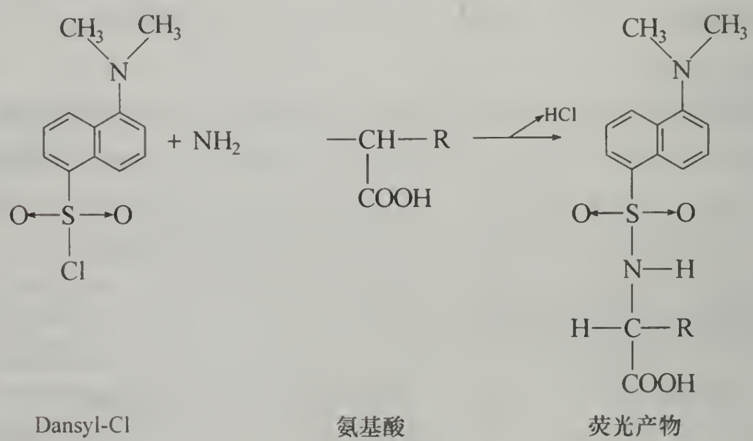
(4) PTC-AA 分析法

属柱前衍生法，源于 Edman 降解法测定蛋白质的一级结构。异硫氰酸苯酯(PITC)能在碱性条件下和氨基酸反应，生成 PTC-AA，此法分析灵敏度和荧光胺、OPA 的荧光法相同，优点是能同时检出初级和次级氨基酸，在 254nm 检出，缺点是操作繁琐，水和盐的副反应敏感，要求较高的操作技术。



(5) Dansyl-Cl 法

能与所有氨基酸柱前反应形成高稳定性的荧光产物,但反应耗时。



(6) Dabsyl-Cl 法

灵敏度是 Dansyl-Cl 法 1/5,但由于反应产生有色衍生物,能方便地在 254nm 和 425nm 处用紫外检测。有报道用 DABITC 代替 Dabsyl-Cl,反应产生 DABTH-AA,在酸性环境中显红色,易在 254nm,269nm 和 436nm 检测。

第六节 测定氨基酸组成的实验步骤

目前各大公司如 Beckman,Hewlett Packard,Perkin Elmer,Waters 等都有各自的将自动进样、氨基酸分析和定量报告联为一体的商品化氨基酸自动分析仪出售,下面介绍几种我们实验室测定氨基酸组成的方法。

一、OPA 法

称取 5.0mg OPA 溶于 0.1ml 甲醇中,加入 50μl 巯基乙醇,用 0.2mol/L pH 9.5 的硼酸缓冲液稀释至 1.0ml。通氮气并避光保存,每隔 2d 加入 5μl 巯基乙醇以维持其强度,此 OPA 反应试剂能用 2~3 周。取 10μl 标准氨基酸混合液(50~100pmol)或样品水解液,加 10μl pH9.5 的硼酸缓冲液(内含 2%SDS 和 100pmol α-氨基丁酸),然后加入 10μl OPA 试

剂，混合均匀，反应1min后再加入20μl 0.1mol/L的KH_2PO_4中止反应，并立即取20μl注入RP-HPLC柱进行分析。

仪器为Beckman 344 HPLC仪，反相柱RP-18(250mm×4.6mm)，荧光检测器。

流动相：A：50mmol/L乙酸缓冲液pH6.8：甲醇：THF=80：19：1

B：50mmol/L乙酸缓冲液pH6.8：甲醇=20：80

梯度条件是：5min，10%B；10min，15%B；20min，10%B；25min，15%B；35min，50%B；50min，90%B；55min，100%B。

我们实验室曾报道过用乙醇代替剧毒的乙腈或甲醇作为有机相来定量分析氨基酸，灵敏度可达10pmol[19]。

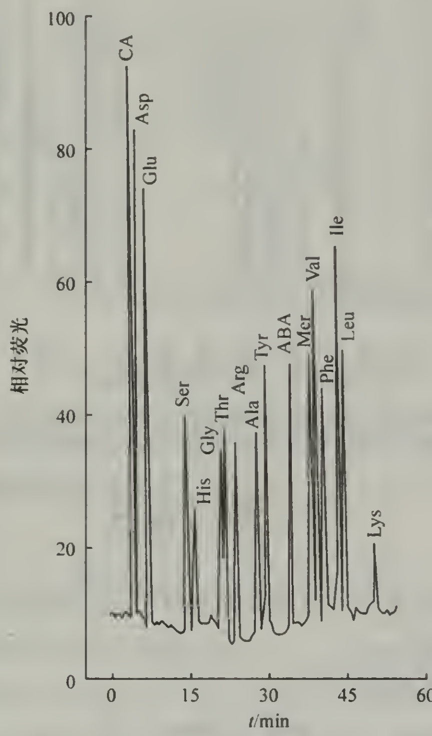


图3.2 OPA法柱前衍生法分析氨基酸

各种氨基酸衍生物均为40 pmol

二、PTC-AA法

氨基酸或样品在Eppendorf管中真空干燥后溶解在50μl耦联试剂(乙醇：三乙胺：水=7：1：1)中。此溶液在旋转真空干燥后再次溶解于50μl耦联试剂中，然后加入2μl PITC，反应混合物在25℃下保温15min，再次旋转真空干燥，真空度(1.3～13Pa)。如此生成的PTC-AA溶解在适量流动相A中，取一部分上柱定量分析[20]。

仪器为Beckman Gold System HPLC仪，反相柱RP-18(250mm×4.6mm)或PTH-柱，254nm检测。

流动相：A：50mmol/L乙酸钠，pH 6.8

B:50mmol/L 乙酸钠,pH6.8:乙腈=50:50

HPLC 柱采用水浴夹套 40℃恒温,并用 5%流动相 B 平衡。梯度条件是:0min,5%B;5min,5%~15%B;10min,15%B;30min,15%~60%B;35min,60%B;40min,60%~5%B。流速是 1ml/min。

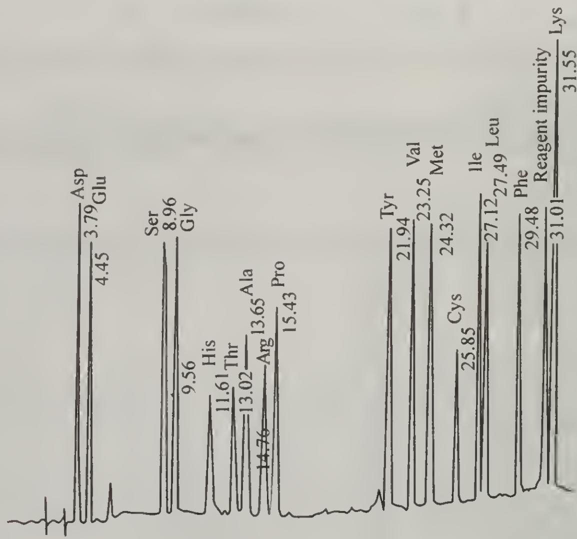


图 3.3 PTC-氨基酸图谱

各种 PTC-AA 100 pmole,254nm 检出,0.010 AUFS

三、茚三酮法

水解后氨基酸样品溶解于 Na-S 样品稀释液中,取 100μl 上样分析,实际进样量 50μl(100pmol~5nmol,均呈线性)。

仪器:Beckman 6300 氨基酸分析仪。

流动相:Na-E 0.00min　　温度:0.00min 48℃

Na-D 27.80min　　11.8min 65℃

Na-F 39.00min　　32.5min 77℃

Na-R 76.00min　　50.0min 48℃

Na-E 77.00~90.00min

流速:　缓冲液 14ml/h

茚三酮 7ml/h

图 3.4 中除常见的 18 种氨基酸(Asn 和 Gln 转变成相应的 Asp 和 Glu)外,还有碘代乙酸还原 Cys 形成的羧甲基半胱氨酸(CM—Cys),过甲酸氧化 Cys 形成的磺基丙氨酸(cysteic acid,CA)。在此系统中,我们观察到 Thr—PO_3,Ser—PO_3,Tyr—PO_3 在溶剂前沿出现,难以分离。这是由于亲水性太强的关系,用毛细管电泳则可将这三种磷酸化的氨基酸分开(参见第十一章)。

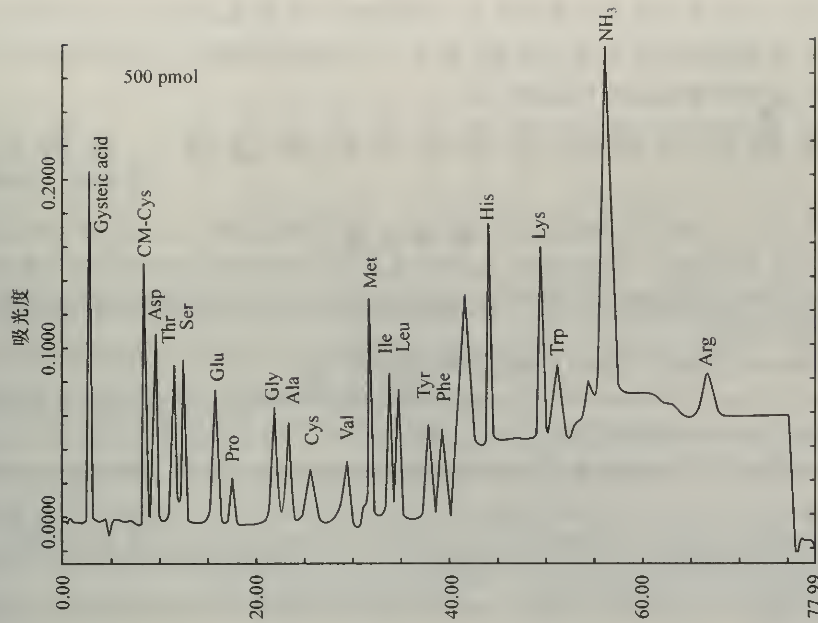


图 3.4　茚三酮柱后法分析氨基酸

四、氨基酸直接分析法

最近美国公司推出的氨基酸直接分析法是基于阴离子交换分离,积分脉冲安培检测。流动相由氢氧化钠和乙酸钠溶液组成[21]。

第七节　结　　语

影响氨基酸分析准确性和重复性的因素有三个:一是样品的水解。水解时间、水解温度、真空度、样品是否被氧化等,都影响实验的结果;二是氨基酸的有效的衍生情况;三是氨基酸色层分离好坏。不同样品采用不同的分析方法,生理样品含有 30～50 种氨基酸,用离子交换色谱柱后衍生法比较合适,天然氨基酸、合成多肽样品的氨基酸测定多采用反相色谱柱前衍生分析方法。食品和饲料中氨基酸分析采用上述两种方法均可。因为反相

表 3.2　一些分析方法的比较

试剂	反应时间	灵敏度	衍生物稳定性	次级氨基酸检测	备注
茚三酮	1～2min	＞100pmol	稳定	能	经典的氨基酸分析法
荧光胺	少于 1s	100pmol(FLD)	稳定	不能	适用于肽类分析
OPA	1min	150fmol(FLD)	不稳定	不能	反应迅速,灵敏度高
PITC	5min	5pmol	较稳定	能	需要样品比较纯净
Dansyl-Cl	60min	5pmol(FLD)	稳定	能	与碱性氨基酸反应产生多种衍生物
Dabsyl-Cl	60min	25pmol	稳定	能	与氨反应会干扰甲硫氨酸的分析

色谱柱前衍生法具有快速和灵敏的特点，在药物和手性氨基酸的分析中同样也获得广泛的应用。几种氨基酸分析法各有利弊(表 3.2)，实验者应根据自己的实际情况，如精确性、方便性、经济性等来选择合适的方法。

(邵晓霞　张丽华)

参考文献

1 Spackman DH, Stein WH, Moore S. Automatic recording apparatus for use in chromatography of amino acids. Anal Chem, 1958, 30: 1190

2 Ozols J. Amino acid analysis. Methods Enzymol, 1990, 182: 587～626

3 Woodward C, Gilman LB, Engelhart WG. An evaluation of microwave heating for the vapor phase hydrolysis of proteins, Int Lab, 1990, September, 40～46

4 Meltzer N M, Tous G I, Gruber S *et al*. Gas-phase hydrolysis of proteins and peptides. Anal Biochem, 1987, 160: 356～361

5 Ng L T, Pascaud A, Pascaud M. Hydrochloric acid hydrolysis of proteins and determination of tryptophan by reversed-phase high performance liquid chromatography. Anal Biochem, 1987, 167: 47～52

6 Bozzini M, Bello R, Cagle N *et al*. Tryptophan recovery from autohydrolyzed samples using dodecanothiol. ABI Research News, Feb. 1991

7 Maeda K. Scheffler JJ. Tsugita A. Tryptophan micro-scale determination by rapid hydrolysis. Hoppe Seylers Z Physiol Chem, 1984, 365: 1183～1185

8 Inglis A S. Single hydrolysis method for all amino acids. Methods Enzymol, 1983, 91: 6～36

9 Hugli TE, Moore S. Determination of the tryptophan content of proteins by ion-exchange chromatography of alkaline hydrolysates. J Biol Chem, 1972, 247: 828～834

10 Delhaye S, Landry J. High-performance liquid chromatography and ultraviolet spectrophotometry for quantitation of tryptophan in barytic hydrolysates. Anal Biochem, 1986, 159: 75～178

11 Edelhoch H. Spectroscopic determinatio of tryptophan and tyrosine in proteins. Biochemistry, 1967, 6: 948～954

12 Greg B, Nice EC, Simpson RJ. Use of scanning diode array detector with reversed-phase microbore columns for the real-time spectral analysis of aromatic acids in peptides and proteins at the submicrogram level. J Chromatogr, 1986, 352: 359～368

13 Hirs CHW. Performic acid oxidation. Methods Enzymol, 1967, 11: 197～199

14 Barkholt V, Jensen AL. Amino acid analysis: determination of cysteine plus half-cystine in proteins after hydrochloric acid hydrolysis with a disulfide compound as additive. Anal Biochem, 1989, 177: 318～322

15 Hoogerheide JG, Campbell CM. Determination of cysteine plus half-cystine in protein and peptide hydrolysates: use of dithiodiglycolic acid and phenylisothiocyanate derivatization. Anal Biochem, 1992, 201: 146～151

16 Szokan G. Pre-column derivatization in HPLC of amino acids and peptides. J Liquid Chromatogr, 1982, 5: 1493～1498

17 Jones BN, *et al*, Amino acid analysis and enzyme sequence determination of protein by an improved O-phthaldialdehyde precolumn labeling procedure. J Liquid Chromatog, 1981, 4: 565

18 Seitz HG, Verhoef M. Automated amino acid analysis using combined OPA and FMOC precolumn derivatization. LC-GC International, 1992, 5: 44～49

19 夏其昌，徐松琴，蔡嘉坤等. 应用乙醇系统的高效液相色层(HPLC)可在 10 微微克分子水平定量分析氨基酸. 生物化学与生物物理学报，1985，17: 748～750

20 夏其昌，吴高德. 反相色层法微量分析 PTC-氨基酸. 生物化学与生物物理学报，1989，21: 465～469

21 牟世芬，丁永胜. 氨基酸直接分析法原理及应用. 现代科学仪器，2001，(2): 24～27

第四章　蛋白质和多肽的氨基酸序列测定

蛋白质和多肽是由20种氨基酸按照一定的顺序通过肽键连接成一长链,然后通过链内、链间的离子键、疏水作用等多种作用力进行折叠、卷曲形成一定的构象并发挥其独特作用。氨基酸的排列顺序即蛋白质的一级结构决定了蛋白质的高级结构及功能。因此,测定蛋白质的氨基酸序列是进行蛋白质结构功能研究中不可缺少的部分。另外,蛋白质一级结构的信息也有助于对其基因结构的研究。

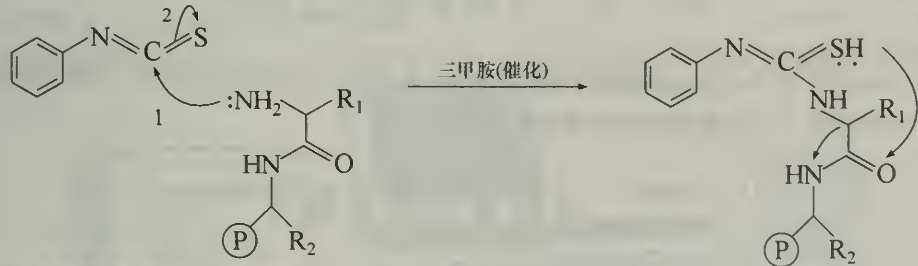
目前,进行蛋白质序列测定有两个关键步骤,即:①氨基酸残基一个一个依次从蛋白质和多肽的末端切割下来;②正确鉴定每次切割下来的氨基酸残基。其中第一步可采用化学法或酶法进行裂解。由于化学法较之酶法具有裂解率高、易于自动化、耗费少等诸多优点,被蛋白质化学实验室普遍采用,可以从蛋白质和多肽的N端进行裂解,也可以从C端进行分析。第二步通常是在氨基酸残基上衍生了一个生色基团,通过高效液相色谱进行分离分析。

本章将就蛋白质和多肽的N端、C端氨基酸序列分析的原理和方法、进行序列分析前蛋白质和多肽样品的准备作一介绍。

第一节　N端分析原理

一、耦　　联

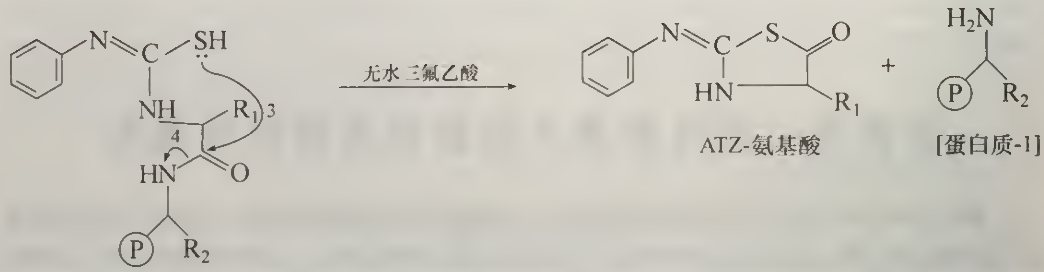
蛋白质和多肽的自由α-氨基经与异硫氰酸苯酯(PITC)试剂耦联后,与其紧邻的第二个残基的键合力大大减弱,很容易断裂。



异硫氰酸苯酯　　蛋白质　　　　　　PTC-蛋白质

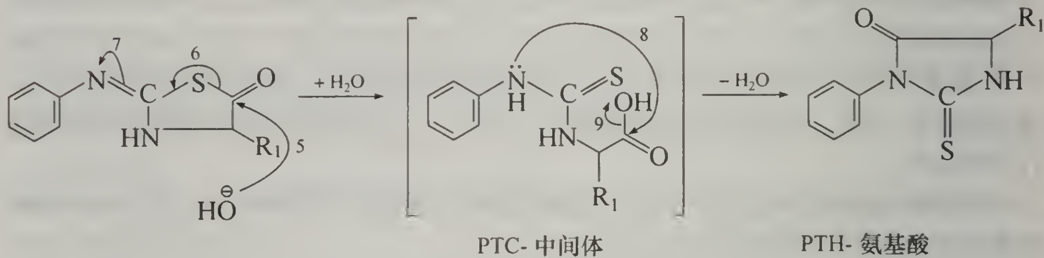
二、裂　　解

在无水条件下酸裂解,仅仅切下反应的那个残基。紧挨的第二个氨基酸残基暴露出自由的α-氨基,又可与PITC进行耦联反应。



三、转　　化

ATZ-氨基酸不稳定,在三氟乙酸水溶液(25%)中可转化成稳定的 PTH-氨基酸。



四、PTH-氨基酸的鉴定

可以通过薄层层析、气相色谱、高效液相色谱及质谱等各种手段进行分析。近年来由于高效液相色谱技术具有分析速度快、灵敏度高、定性定量准确等诸多优点,并且仪器自身结构也日臻完善,被广泛应用于PTH-氨基酸的鉴定。通常选用C-18反相柱进行分离。

第二节　N端蛋白质序列仪

自20世纪70年代初自动序列仪诞生以来,经过了一系列更新换代,仪器日趋灵敏、快速,下面就其中几种代表性的仪器作一简介。

一、液相旋转杯序列仪

这是最初的自动化仪器。整个仪器的"核心"部分是一个反应杯,它由一个马达带动并有调速控制。蛋白质或多肽样品经溶解后,通过一根导管注入反应杯中,样品随着旋转的离心力被均匀地涂在杯壁上,形成一层薄膜,其厚薄可由转速控制。反应的试剂和溶剂分别通过导管进入反应杯,与杯内薄层上的样品的N端自由氨基进行Edman反应;降解下来的ATZ氨基酸衍生物通过另一导管引出并收集,再将ATZ氨基酸转换成PTH氨基酸后进行鉴定,依次循环分析。此仪器由于消耗的样品量多,目前在蛋白质化学实验室已很少使用。

二、固相序列分析仪

由于一些小肽或疏水性多肽，不易吸附在一般支持物上(如旋转杯)，只得将它们共价固定在惰性载体后进行 Edman 化学降解反应。固相测序的优点是样品共价结合后，不会在有机溶剂洗涤、抽提等过程中丢失，因而可以增加循环次数。其缺点是共价结合不外乎通过氨基或羧基与载体上的活性基团作用，蛋白质和多肽中的谷氨酸、天冬氨酸、赖氨酸等含有除 α-氨基或 α-羧基以外氨基或羧基的氨基酸残基将结合在载体上，不能正确鉴定。

三、气相序列仪

自动化蛋白质序列仪刚问世时，需要样品量为 1～2mg。1978 年 Tarr 等将一种四级铵盐聚合物“polybrene”引入了蛋白质、多肽的序列测定，它能和肽链中的极性基团作用(非共价键合)而将蛋白质和多肽有效地固定在玻璃纤维支持膜上，防止了萃取时样品的损失。为了适应分子生物学对蛋白质微量分析的要求，20 世纪 80 年代初，Hewick 及 Hunkpillar 等人改进了自动化蛋白质序列仪中的仪器结构，它用弹筒型玻璃反应室(内部体积约 0.05ml)代替了液相序列仪中的旋转玻璃杯，用 polybrene 固定样品，Edamn 降解反应中有的试剂如三甲胺以气体方式输送，其他试剂以流动或液相脉冲方式输送。此仪器较之以前几种序列仪具有以下明显优点：

①灵敏度高，耗费样品量少，通常仅需 50～100pmol；

②试剂和溶剂消耗少，大约是液相序列仪的 1/10，降低了费用；

③缩短了每步降解循环时间，提高了分析效率。

20 世纪 80 年代末又对气相序列仪进行了部分改进，即将原来三氟乙酸以气体方式输送改为液相脉冲方式，称为脉冲液相序列仪，以适应不同样品的分析需要。另外，由于载有活性基团的功能性 PVDF 膜的问世，蛋白质和多肽可以共价固定在此类膜上，直接在上述两种序列仪上进行固相序列分析。

在这两种序列仪中，Edman 降解后的 PTH-氨基酸衍生物可及时直接从转化腔中自

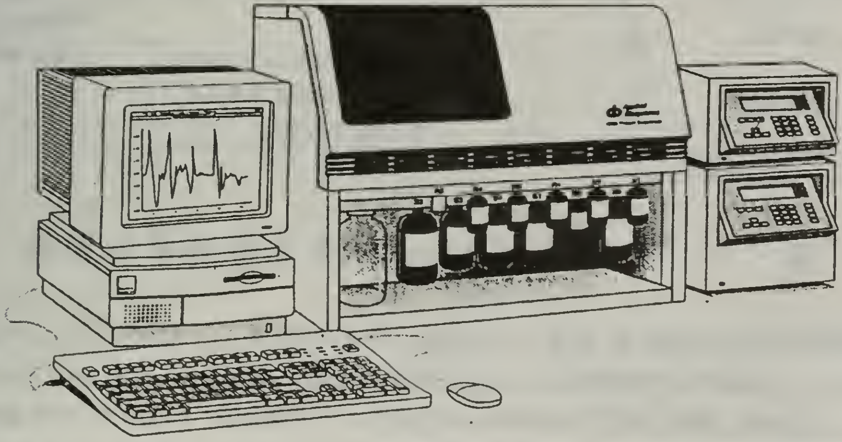


图 4.1　序列仪结构示意图

动输入高效液相色谱仪中进行 PTH-氨基酸衍生物的定性及定量分析,这样不仅快速可靠,而且防止了样品的损失。这两台仪器统一用电脑控制,包括化学反应和分析鉴定以及数据的自动记录和处理(图 4.1)。图 4.2 为标准 PTH-氨基酸的色谱分离图。由图可见,欲在 20min 内将各组分较好分离,需选择合适的色谱条件,表 4.1 列出了各种条件的改变导致个别组分位置的迁移及图谱的改善。

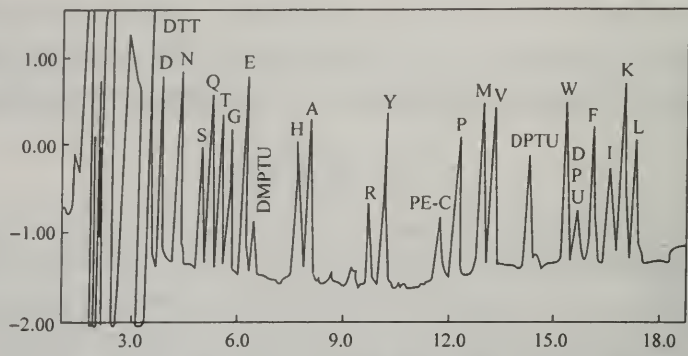


图 4.2 PTH-氨基酸色谱图

表 4.1 PTH-氨基酸分离条件的调整

改变下列条件	结果		改变下列条件	结果		
增加起始 B 液百分比	←	→	增加 A 缓冲液浓度	←	←	←
	E	DMPTU		H	R	PE-Cys
	←	→	降低 A 缓冲液浓度	→	→	→
	Q	T		H	R	PE-Cys
降低起始 B 液百分比	←	→	降低 A 液 pH 值	→	→	
	S	Q		D	E	
	→	→	增加 A 液 pH 值	←	←	
	DTT	D		D	E	
柱温降低 3℃	→	→	需更换 A 液(因 THF 浓度	←	→	
	DTT	D	降低)	Q	T	
柱温升高 3℃	←	→		←	→	
	M	V		P	M	
降低最终 B 液百分比	←	→	用酸洗泵	可使基线平坦		
	I	K	1L A液中加 0.5ml 12.5%TMA	H 和 R 峰尖锐		
增加最终 B 液百分比	←	→	(可能需用 A 液稀释来调节	←	←	←
	K	L	H 和 R 的位置)	H	R	PE-Cys

注:←、→ 为色谱峰迁移方向。

第三节 影响 N 端 Edman 反应裂解率的因素

在测序中往往可以发现,即使 N 端很均一的蛋白质(第一个循环出现单个氨基酸残基),经过十几个循环后背景明显不及第一个残基清晰。这是由于 Edman 反应是一个化学过程,在其耦联、裂解、转化等步骤都有可能发生一些副反应,从而影响最终裂解效率。目前最佳产率为 95%～98%。因此,经过 50 次循环后,PTH-氨基酸分析谱中将出现较

多杂峰，影响正确辨认，也就是说对于一般蛋白质最多能分析至 N 端第 50 个氨基酸左右，欲知全顺序，则需将蛋白质降解成一系列肽段分析后再拼接。

对于有些蛋白质由于含有较多的对 Edman 反应敏感的残基或肽键，裂解效率将更低。表 4.2 列出了在蛋白质和多肽中容易发生断裂的一些肽键。

表 4.2　容易发生断裂的肽键

肽键	断裂百分比/循环
↓ X—Ser—X	0.1～0.15
↓ X—Ser—Thr	0.35
↓ X—Thr—X	0.25～0.5
↓ X—Asp—X	<0.02

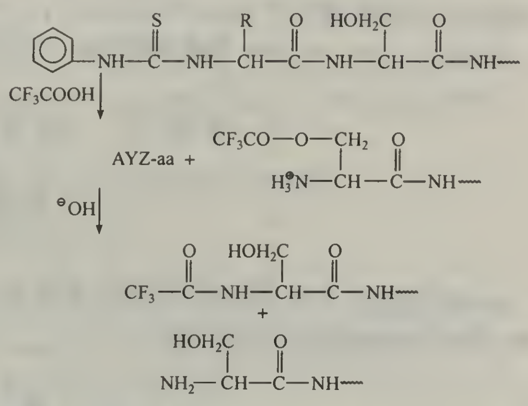


图 4.3　三氟乙酸造成的 Ser 和 Thr 的 N 端封闭

有时循环至 Ser 和 Thr 时产率突然降低，这是因为在进行这两个氨基酸残基分析前一个循环中，三氟乙酸除起裂解作用外，可能与 Ser 和 Thr 上的羟基反应，继而部分封闭 Ser 和 Thr 的 N 端 α-氨基，见图 4.3。

另外，丝氨酸和苏氨酸的 PTH 衍生物也会部分转化成其他产物，造成产率的降低，见图 4.4。

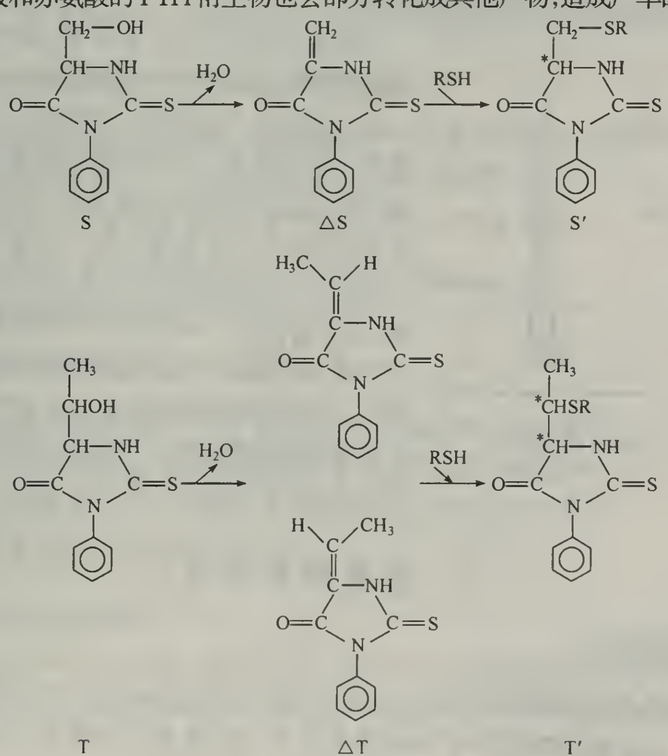


图 4.4　PTH-Ser 和 PTH-Thr 转化成副产物

耦联试剂 PITC 本身也将发生下列副反应,形成一些“杂质”。

测序完成后,电脑将每个循环的产率处理,得出起始产率(initial yield)和重复产率(repetitive yield),其中通过起始产率可以估计蛋白质的真实含量,有时可以推测 N 端是否封闭(和氨基酸组成分析测得的含量相差甚远),而重复产率则表明机器是否正常运转。但是对于纯度不够的样品,上面两项均不具备实际意义。另外,如果蛋白质或多肽中含有过多的 Ser、Thr、Glu、Trp 等氨基酸残基,也将降低这两个数值。

第四节　测序前样品处理

一、纯度鉴定

宜采用多种互补有效的手段鉴定样品的纯度,如 SDS-PAGE、反相 HPLC、毛细管电泳、阴离子或阳离子的 FPLC 等。

二、脱　　盐

可采用多种有效的脱盐方法如反相 HPLC、凝胶过滤、透析、超滤等,但需注意在进行脱盐过程中除需考虑是否除盐完全外,所用试剂、仪器乃至操作规范也必须是测序级的,应尽量避免引入新的杂质。下面介绍一种简便的 ProSpin 装置,尤其适合于普通实验室对蛋白含量少的样品进行序列分析前脱盐。

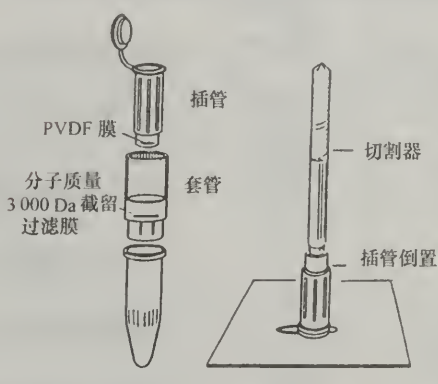


图 4.5　ProSpin 示意图

图 4.5 为 ProSpin 示意图,其中在插管的底部黏附了 ProBlott PVDF 膜(有关 ProBlott 膜见电泳一章),插管插入套管后,此膜与分子质量 3 000Da 截留过滤膜接触,再在套管的底部连接一离心管,通过离心除去样品中的缓冲盐、去垢剂及其他小分子杂质,也可以在套管下放一海绵(接触滤膜),将上述小分子物质随水分一起吸去。然后将插管倒置,用切割器取下吸附有蛋白质的膜片(ProSpin 和切割器由 Perkin-Elmer 公司提供,并附详细操作步骤)。对于特别稀少的样品可将切下的膜片分割成几部分,分别进行测序、氨基酸组成分析、肽谱分析等工作,其中每部分约需 50pmol。

三、巯基修饰方法

1. 丙烯酰胺修饰[1]

第 1 步:还原:蛋白质溶解于 10～15μl 0.2mol/L Tris,pH8.4,内含 100mmol/L DTT 缓冲液中,加入 SDS 至最终浓度为 1%,将样品于 70℃ 水浴中保温 20～30min,然后用 4 倍体积的重蒸水稀释。

第 2 步：烷基化：加入 6mol/L 丙烯酰胺浓溶液至最终浓度为 2mol/L，通入氩气（或高纯氮气）后于 37℃ 避光保温 30～60min。

第 3 步：脱盐：加入甲醇至最终浓度为 10%，再用 Perkin-Elmer 公司的 ProSpin 管进行脱盐。

具体操作：先将 10μl 甲醇润湿 ProSpin 上的 ProBlott 膜，然后加入反应溶液离心至膜上面无液体，再加入 50μl 20% 甲醇，同上进行离心，将膜片切割下后放入 1.5ml 离心管中，用 0.1%TFA 清洗后，再用水漂洗，待其自然晾干后（或用氮气吹干），即可进行序列分析。

注意：第 3 步也可采用其他有效的脱盐法。但上述方法相对比较简便。

2. 4-乙烯吡啶修饰[2]

第 1 步：还原：蛋白质溶于 40μl 100mmol/L Tris，pH 8.4，内含 6mol/L 盐酸胍缓冲液中，加入 1mol/L DTT 浓溶液到最终浓度为 20mmol/L，充入氩气后于室温放置 1～2h。

第 2 步：烷基化：加入 2μl 4-乙烯吡啶后充分混合，同上保温 1～2h。

第 3 步：脱盐：加入重蒸水 1:1 稀释蛋白质溶液后，同前述用 ProSpin（或其他方法）脱盐。

四、去除 N 端封闭的基团的方法

如果怀疑蛋白质或多肽的 N 端被封闭，可按下述方法去除一些 N 端封闭的基团[3]。

1. 去除 *N*-乙酰丝氨酸和 *N*-乙酰苏氨酸残基

第 1 步：蛋白质吸附在 PVDF 膜上，经蛋白质 N 端序列分析几个循环后，未能检测出 N 端氨基酸残基，将膜条置于 1.5ml 的塑料离心管中。

第 2 步：加入 50μl 三氟乙酸液体后密封管口并在 40℃ 保温 1h。

第 3 步：打开管盖，在通风橱中使 TFA 挥发完全。

第 4 步：将此膜条再放回序列仪中进行分析。

2. 去除 N 端甲酰甲硫氨酸中的甲酰基

第 1 步：将吸附了蛋白质的 PVDF 膜条置于 1.5ml 塑料离心管中。

第 2 步：加入 30μl 0.6mol/L HCl 后密封管口并于 25℃ 保温 24h。

第 3 步：打开管盖，吸去 HCl 溶液，再将膜条真空干燥或氮气吹干。

第 4 步：将膜条放回序列仪中继续分析。

3. 去除 N 端的焦谷氨酸

第 1 步：将吸附了蛋白质的 PVDF 膜条置于 1.5ml 的塑料离心管中，加入 1.2ml 0.5%（*m*/*V*）PVP-40 封闭膜条，使其不再吸附其他蛋白质（PVP-40 溶于 100mmol/L 冰醋酸中），于 37℃ 保温 30min，然后用超纯水清洗膜条至少 5 遍，接着再用 0.5ml 酶解缓冲液（下一步将用）漂洗一遍。

第 2 步：加入 100μl 左右焦谷氨酸水解酶（5μg 酶溶于 100μl 50mmol/L，pH7.0 磷酸钠及 10mmol/L DTT 缓冲液中），37℃ 保温 5～10h。

第 3 步：用超纯水漂洗膜条，自然吹干后放回序列仪中。

4．去除 *N*－乙酰基封闭的氨基酸残基

第 1 步：将吸附了蛋白质的 PVDF 膜条置于 1.5ml 的塑料离心管中。

第 2 步：用 0.5% PVP-40（溶于 100mmol/L 冰醋酸）封闭 PVDF 膜以防吸附蛋白酶。

第 3 步：将膜用超纯水充分清洗后，加入 5～10μg 胰蛋白酶（溶于 100μl 0.1mol/L 碳酸氢铵，pH8.0，内含 10% 乙腈），于 37℃ 酶解 24h。

第 4 步：将离心管中的溶液移至另一干净的离心管中，用 100μl 超纯水清洗膜条后将此清洗溶液与酶解后的溶液合并，然后将此含有酶解多肽的溶液冷冻干燥。

第 5 步：加入 100μl 50% 吡啶及 10μl 异硫氰酸苯酯（PITC），通入氮气 20s，然后于 60℃ 保温 1h。此步骤中 PITC 可以和酶解后各多肽的 N 端 α－氨基反应（除 N 端被乙酰基封闭的多肽）。

第 6 步：加入 50μl 苯/乙酸乙酯（1∶1 的体积比）混合溶剂，充分振荡后以 3 000g 离心 1min。

第 7 步：吸去上层溶液（内含第 6 步中的过量试剂及反应副产物）。再按第 6 步对下层溶液抽提 3 次，然后真空抽干下层液。

第 8 步：加入 100μl 过甲酸（9 份甲酸与 1 份过氧化氢混合后在室温放置 1h），于 0℃ 保温 1h，以封闭经过上述几步反应的多肽的 N 端。

第 9 步：真空抽干，再用水溶解，然后再将样品抽干。

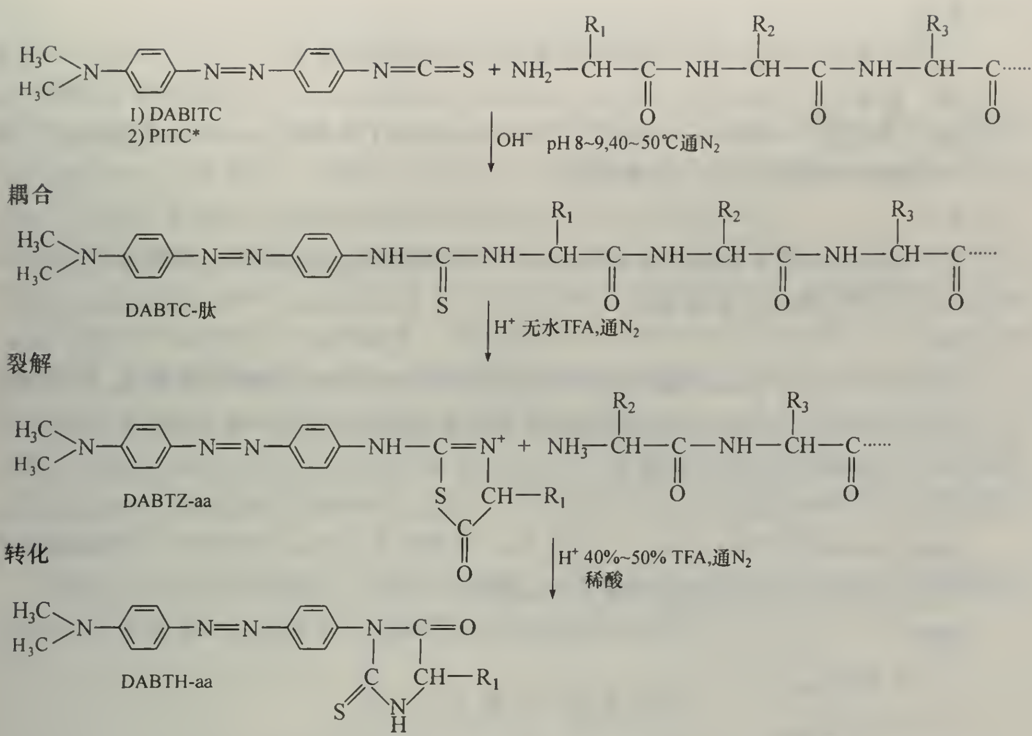
第 10 步：将多肽用 100μl 0.2mmol/L，pH 7.2 磷酸缓冲液（内含 1mmol/L DTT）溶解，加入 0.05 单位的 *N*－乙酰氨基酸去除酶（AARE），于 37℃ 保温 12h。

第 11 步：将反应混合物加样至已预处理过的测序用玻璃纤维支持膜上，并放入序列仪中分析。

第五节　手工 N 端测序

虽然自动化测序方便、快速，但并不能完全排除手工测序的实用性，尤其对于一些实验室尚不具备测序条件，而且样品量足够，只需知道 N 端个别残基以确定是否均一，手工测序不失为一种简便有效的方法，下面简单介绍一种手工测序方法——DABITC/PITC 双耦合法[4]。

DABITC/PITC 双耦合微量顺序测定法，采用有色试剂 DABITC 与多肽或蛋白质的氨基作第一次耦合，因该耦合反应产率只有 20%～50%，余下没有与 DABITC 耦合的氨基再用 PITC 作第二次耦合，由于 PITC 与多肽或蛋白质的耦合反应几乎是定量的，这样就保证了每个残基都反应完全，然后再裂解、转化、鉴定有色的 DABTH－氨基酸。失去一个残基后的肽或蛋白质再耦合、裂解、转化，依次循环测定多肽和蛋白质序列，其反应原理如下：



DABITC:4-*N*,*N*-二甲胺基偶氮苯-4′-异硫氰酸酯　(4-*N*,*N*-dimethylaminoazobenzene-4′-isothiocyanate)

DABTC:4-*N*,*N*-二甲胺基偶氮苯-4′-硫代氨甲酰基　(4-*N*,*N*-dimethylaminoazobenzene-4′-thiocarbamoyl)

DABTZ-aa:4-*N*,*N*-二甲胺基偶氮苯-4′-噻唑啉酮-氨基酸　(4-*N*,*N*-dimethylaminoazobenzene-4′-thiozolinone-aa)

DABTH-aa:4-*N*,*N*-二甲胺基偶氮苯-4′-乙内酰硫脲-氨基酸

(4-*N*,*N*-dimethylaminoazobenzene-4′-thiohydantoin-aa)

PITC 耦合,裂解,转化产生的 PTH-氨基酸不作鉴定

(1) 材料

DABITC(Fluka 公司出品或自己合成),用沸丙酮重结晶。n-正庚烷、乙酸乙酯、n-乙酸丁酯都按顺序级要求处理。吡啶先用 KOH(10g/L)回流,蒸出,再用茚三酮(1g/L)回流,蒸出,最后再用 KOH(10g/L)回流,再蒸出。异硫氰酸苯酯(PITC)在减压下通氮气重蒸。三氟乙酸用 CrO_3 处理重蒸再用 $CaSO_4$ 干燥重蒸。超滤水。

聚酰胺薄膜层析溶剂:A. Ⅰ相 33%乙酸;Ⅱ相甲苯:n-乙烷:乙酸(2:1:1,体积比)。B. 鉴定 Ile,Leu,膜 7.5cm×4cm,10%甲酸:乙醇(10:9,体积比)。

高效硅胶板(HPTLC Merck)鉴定 Ile,Leu,氯仿:甲醇(50:2,体积比)。

(2) 耦合

将 2~5nmol 样品放在尖底带塞小试管中(Edman 管),加 80μl 50%吡啶溶液,加 40μl DABITC 吡啶溶液(0.6mg/260μl),通高纯氮(99.99%),盖好塞子,在 50℃加热 45min,然后再加 10μl 异硫氰酸苯酯,通氮气,用混合仪使 PITC 与反应溶液很好混合,在 50℃加热 30min,反应混合物用 4×200μl n-正庚烷:乙酸乙酯(2:1)(如疏水性肽用 3:1)提取,除去有机层,水层在真空干燥器中干燥。

(3) 裂解

干燥后水层产物加 50μl 三氟乙酸,通入氮气温和搅拌混合后,在 50℃ 加热 15min,真空干燥。如是 Pro—Pro;Pro--Ile;Pro—Val;Ile—Ile;或是 Val—Val 则需要重复裂解 1~2 次,裂解干燥后的产物加 30μl 双蒸水,用 2×50μl 乙酸丁酯提取 DABTZ-氨基酸,并干燥之。水层真空干燥后,作下一步循环用。

(4) 转化

将干燥后的乙酸丁酯提取物加 30μl 40%~50%三氟乙酸水溶液,通氮气,在 50℃ 加热 50min,真空干燥后可作鉴定。

(5) 鉴定

干燥物加 10~20μl 乙醇溶解,取 1/2μl 点到 3cm×3cm 聚酰胺薄膜上,然后再加 1/2μl标记化合物,按上面 A 溶剂系统展开,其标准层析图谱见图 4.6。如鉴定 Ile、Leu,则可用溶剂 B 系统或硅胶板溶剂系统。

(6) 标记化合物的制备

500μl 50%吡啶中加入 30μl 二乙胺,30μl 乙醇胺及 330μg DABITC(1 000nmol),在 55℃ 加热 1h,真空干燥,用时取少量溶于乙醇中。

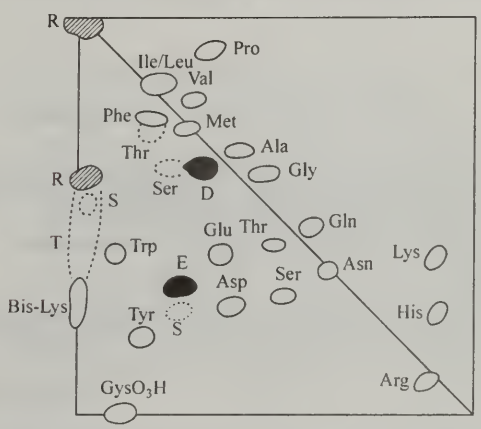


图 4.6 DABTH-氨基酸在聚酰胺薄膜上的层析图谱

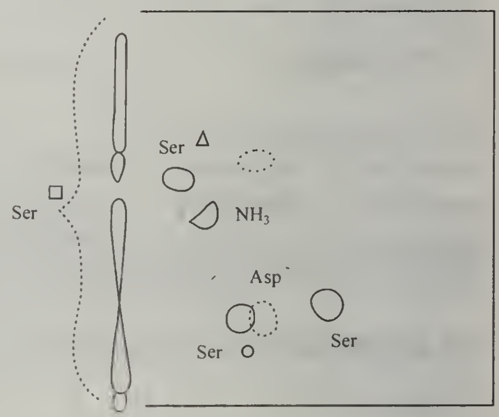


图 4.7 四种 Ser 产物在聚酰胺薄膜上的层析图谱

注意

①所用试剂药品都要按顺序级要求处理。DABITC 在吡啶中不稳定,必须新鲜配制。

②疏水性肽在用 2:1 或 3:1 的庚烷和乙酸乙酯(体积比)提取时易产生沉淀或薄膜,悬浮在有机相和水相中间,切勿吸去。

③ Gln、Asn 可能有 8%~15%水解成 Glu、Asp,因此在鉴定 Gln、Asn 时有时也出现很淡的 Glu、Asp 斑点。

④层析溶剂均应新鲜,保证层析的重复性。

⑤ Lys 可出现三个不同颜色的衍生物,α-DABTH-ε-DABTC-Lys(紫色);α-ε-DABTC-Lys(蓝色)和 α-DABTH-ε-PTC-Lys(红色)。

⑥ Ser 和 Thr 副反应较多,一般可以有 3~4 个产物:DABTH-Ser;DABTH-Ser△(失水产物);DABTH-Ser○(失水后再聚合产物);DABTH-Ser□(失水后再加水产物)。

第六节　C端序列分析

虽然建立在Edman化学法上的自动化N端测序技术日趋成熟,但仍有不少问题需借助C端序列分析来解决。C端测序尤其适合于基因重组蛋白是否正确表达的鉴定和大规模生产时的质量控制、N端封闭蛋白的分析以及DNA探针的设计。

由于选择性对C端羧基进行耦联修饰并从C端逐一将氨基酸残基切下较N端测序困难,在过去的近20年里,虽然为数不少的蛋白质化学家致力于C端测序研究[5],但目前仍不能和N端测序技术程度媲美。蛋白质化学工作者曾利用羧肽酶分析C端氨基酸残基[6,7],但是由于不同的羧肽酶对个别氨基酸残基有选择性,使用时仍有一定困难。另有报道[8]将C端选择性标记,分离出C端肽段后,再用Edman法分析,我们实验室[9]曾根据一些蛋白接近C端的个别氨基酸(如Trp,Tyr)有特殊的紫外光谱或二级微分光谱,分离出含有C端氨基酸的酶解肽段,进行N端序列分析,以此鉴定蛋白质的C端是否有缺失残基。

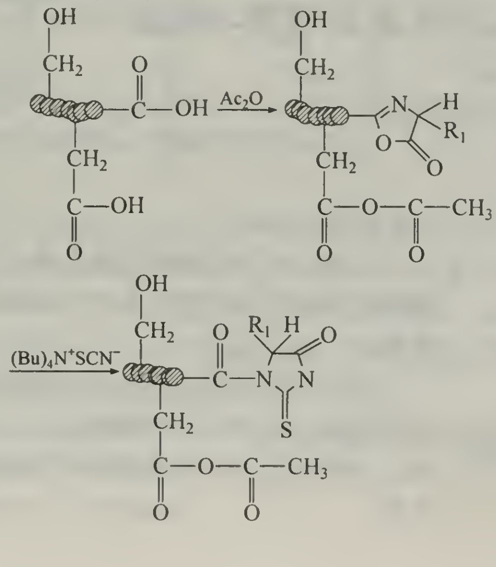
最近,由于对化学法测定C端技术进行了一系列改进,已有了自动化C端序列仪问世[10]。下面介绍一种自动化C端测序方法的原理及应用[11]。

一、C端分析原理

基于与N端Edman降解类似的原理,C端分析采用化学试剂与蛋白质和多肽的α-羧基反应,反应后的C端衍生物被切割下来,通过HPLC系统进行分离分析。一个完整的C端序列分析包括以下几个步骤。

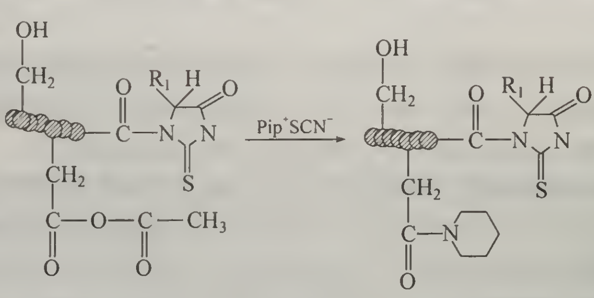
(1) 活化

蛋白质和多肽C端的自由羧基与乙酸酐反应生成混合酸酐,并进一步环化成噁唑酮,而侧链上的羧基形成的混合酸酐难以成环。C端的噁唑酮在硫氰四丁铵的作用下,转化为乙内酰硫脲(thiohydantoin,TH)。



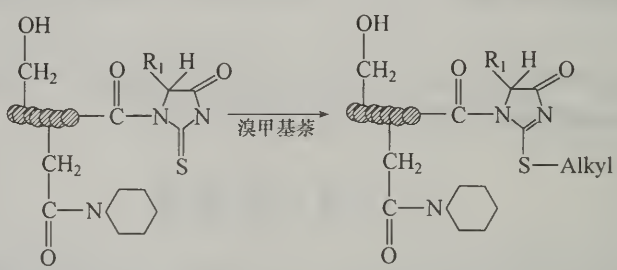
(2) 酰胺化保护侧链羧基

侧链羧基形成混合酸酐后，与硫氰哌啶进一步作用形成酰胺，以保护侧链。



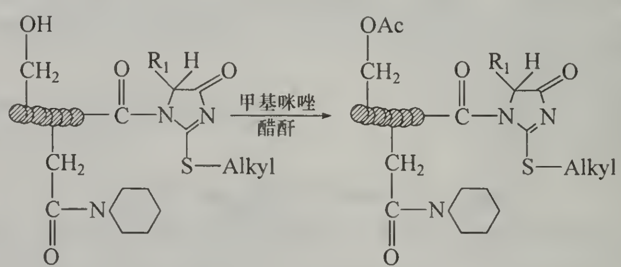
(3) 烷基化

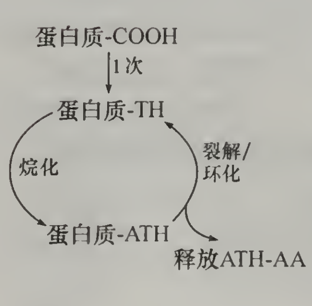
由于C端环化形成的TH衍生物较为稳定而不易被切割，因此产率不高。用溴甲基萘选择性地烷基化修饰硫原子，形成Alkylated-TH (ATH)衍生物，裂解产率将大大提高。



(4) 修饰Thr和Ser的羟基

蛋白质和多肽中的羟基会干扰测序，因此需要修饰。在活化过程中，乙酸酐也会与羟基反应将其乙酰化，但反应不完全，在 N -甲基咪唑(NMI)和乙酸酐的共同作用下，Thr和Ser上的羟基基本被乙酰化。





(5) 裂解和衍生

C端的ATH衍生物在酸性条件下与[NCS]⁻反应而被裂解生成ATH-氨基酸，新的C端自动形成TH，而不必重新活化。

因此，整个测序过程只需在开始时对C端进行一次性活化，并修饰Asp和Glu侧链羧基以及The和Ser的羟基。实际上，循环反应的只有烷基化和裂解两步，其全过程图解如左图。

二、C端蛋白质序列仪

近年来，已有商品化的C端序列仪问世。C端序列仪一般由N端序列仪改装而成，其基本结构与后者相似，两者的不同之处主要在于：

①反应所用的试剂不同，因此两者不可兼容，否则极易堵塞管道；

②在C端序列仪中，所有的化学反应均在弹筒型反应室中进行，切割下来的ATH-AA仅在转化腔内干燥和溶解后即进入HPLC系统分离分析；而在N端测序仪中，ATZ-AA需在转化腔中转化为PTH-AA。参见图4.8。

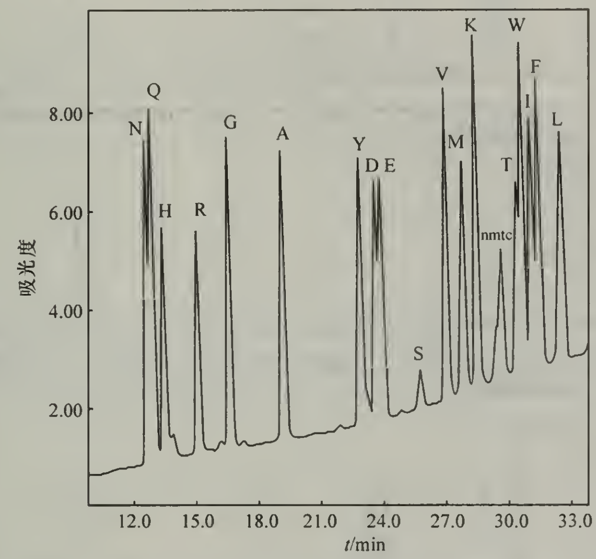


图4.8 ATH-氨基酸色谱图

三、C端测序的样品前处理

C端测序样品的纯度鉴定和巯基修饰方法与N端测序基本相同，可参照前述。

下面介绍C端测序样品的脱盐和游离氨基的衍生。

(1) 脱盐的方法

目前使用最普遍的C端脱盐方法是采用ProSorb装置。ProSorb的结构和使用方法与ProSpin大致相同，不同的是用过滤器代替了ProSpin中的离心管，过滤器可将样品中的小分子物质随水分一起吸收，以达到脱盐和浓缩样品的目的。具体操作步骤如下。

第一步：将插管插入套管上部，过滤器插入套管下部并接触插管底部的PVDF膜。

第二步：用15μl甲醇润湿PVDF膜。

第三步：在插管中加入出50μl 0.1%TFA，待过滤器将其完全吸干。

第四步：将样品溶液加入，体积不少于100μl，吸干。

第五步：用TFA洗膜1～2次，吸干。

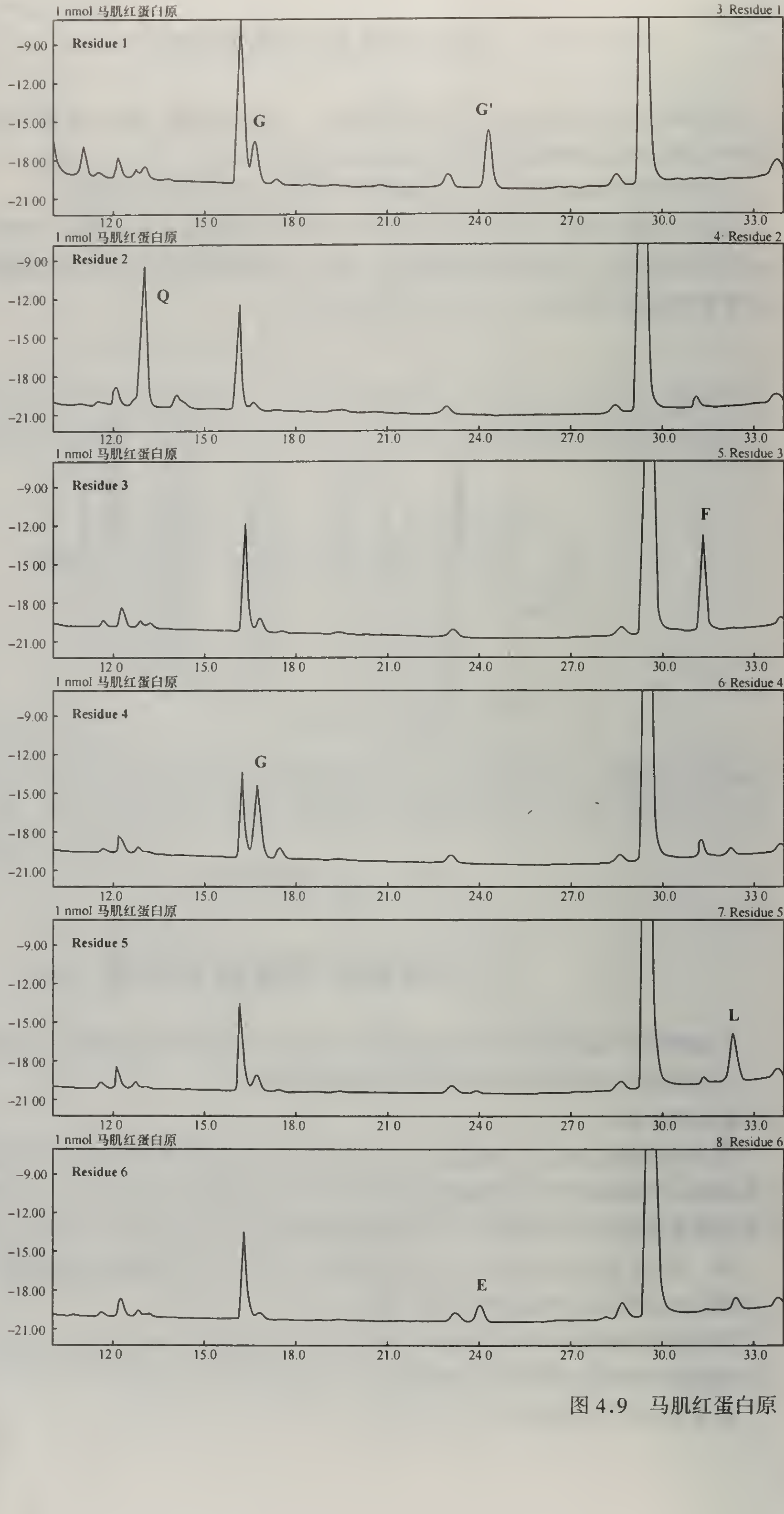
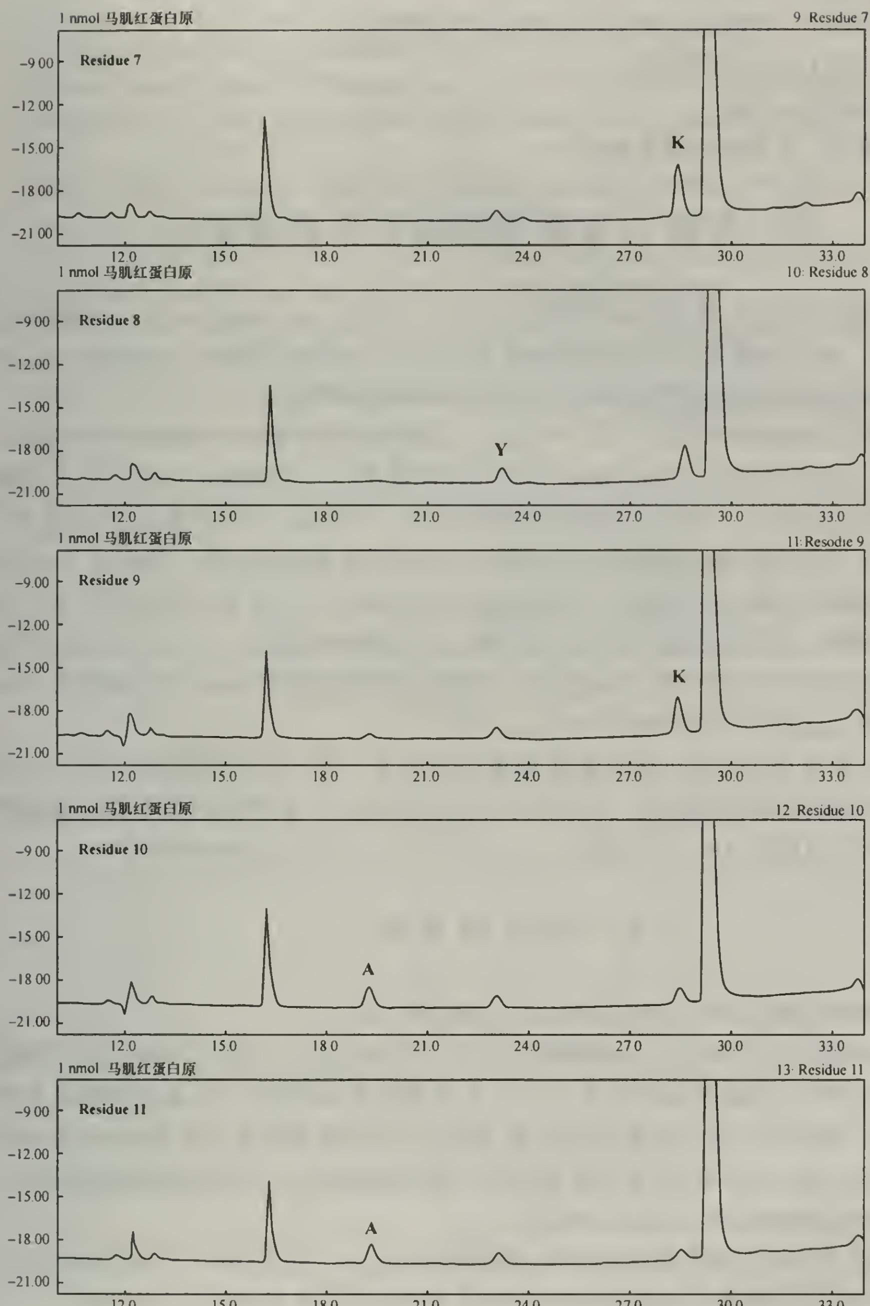

1 nmol 马肌红蛋白原
3 Residue 1
Residue 1
G
G'
1 nmol 马肌红蛋白原
4 Residue 2
Residue 2
Q
1 nmol 马肌红蛋白原
5 Residue 3
Residue 3
F
1 nmol 马肌红蛋白原
6 Residue 4
Residue 4
G
1 nmol 马肌红蛋白原
7 Residue 5
Residue 5
L
1 nmol 马肌红蛋白原
8 Residue 6
Residue 6
E

图 4.9　马肌红蛋白原


1 nmol 马肌红蛋白原
9 Residue 7
Residue 7
K
1 nmol 马肌红蛋白原
10: Residue 8
Residue 8
Y
1 nmol 马肌红蛋白原
11:Resodie 9
Residue 9
K
1 nmol 马肌红蛋白原
12 Residue 10
Residue 10
A
1 nmol 马肌红蛋白原
13 Residue 11
Residue 11
A

的C端测序图谱

第六步：取出插管，置于恒温加热块上，55～60℃加热至少 10min。

(2) ε-氨基的衍生

为了使 Lys-ATH 衍生物获得更好的分离效果，需先将蛋白质或多肽中 Lys 的 ε-氨基与 Phenylisocyanate (PhNCO)反应成衍生物。具体操作步骤如下。

第一步：脱盐后，从恒温加热块上取出插管。

第二步：用 3μl 二异丙基乙胺(DIEA)润湿 PVDF 膜。

第三步：在膜上加 3μl PhNCO。

第四步：加热 5～10min。

第五步：将第二至第四步重复两次。

四、影响 C 端测序反应产率的因素

C 端测序副反应较多，最高初始产率仅为 20%，而对于 Glu，Asp，Ser，Thr 等氨基酸，其产率将更低。如 C 端富含上述几种氨基酸，测序将十分困难。另外，一旦遇到 Pro，由于吡咯环的存在而不能与乙酸酐反应成环，整个测序过程将终止。

到目前为止，C 端测序可测 1～10 个残基，一次测序需要的量比 N 端测序要大得多，至少需 1nmol。此外，因为不同的氨基酸反应产率不同，所以，对图谱的分析也比 N 端测序复杂，需要较多的经验。目前 C 端测序结果最好的一个样品是马肌红蛋白原(图 4.9)，可测 C 端 10 个以上残基，通常将其作为标准样品来判断仪器的稳定性。我们对 TNF 的 C 端测定结果得到 C 端 8 个序列为—VYFGIIAL—COOH。由于需要对样品进行 PIC 预处理和侧链的修饰，有些氨基酸的 ATH 衍生物上会出现修饰基团。如：Asp，Glu 将生成侧链酰胺化的 ATH 衍生物；Ser，Thr 通过 β-消除分别生成脱氨 Ala-ATH 和脱水 Thr-ATH 的衍生物；Lys 的反应终产物为 PIC-Lys-ATH。

另外，由于副反应的存在，有的氨基酸将生成不止一种 ATH 衍生物。如：Arg 的 ATH 衍生物可能被乙酰化或烷化；Tyr-ATH 可被乙酰化；Cys 被修饰后将形成丙烯酰胺化的 Cys-ATH 以及脱氢 Ala-ATH；脱水 Thr-ATH 将产生两个非对映异构体。

五、结语和展望

目前，C 端序列测序技术已初步成熟，并开始发挥作用。

但其反应的效率与 N 端 Edman 降解测序法有一定的差距。因此，目前的主要问题是如何提高反应的产率，特别是如何改善 T，S，D，E 几种氨基酸的测定，以及如何解决 P 终止测序的问题。如果在上述几方面取得进展，将使 C 端可测残基数大大增加，并减少样品所需量。另外，近年来发现，很多活性多肽的 C 端是酰胺化的，如果能检测酰胺化的 C 端残基，将对活性多肽的研究有很大的帮助。

C 端测序技术仍处于研究和发展阶段，虽然该技术还不尽完善，但对于确认表达蛋白的 C 端是否正确，以及判断在表达、纯化过程中是否发生了不必要的加工大有帮助。

（徐来根　屠红旻　曾　嵘）

参考文献

1 Brune D C. Alkylation of cysteine with acrylamide for protein sequence analysis. Anal Biochem, 1992, 207: 285～290

2 Fullmer C S. Identification of cysteine-containing peptides in protein digests by high-performance liquid chromatography. Anal Biochem, 1984, 142: 336～339

3 LeGendre N *et al*. Purification of proteins and peptides by SDS-PAGE. In: Matsudaira P. ed. A pratical guide to protein and peptide purification for microsequencing, 2nd ed. USA: Academic Press, 1993

4 Wittmann-Liebold B, Kimura M. Microsequencing of peptide and proteins with 4-N, N-dimethylamino-azo-benzene-4′-isothiocyanate. In: Walker J M ed. Methods in molecular biology. Vol. I. Proteins, The Humana Press Inc, 1984, 221～241

5 屠红旻,夏其昌.蛋白质和多肽 C 端氨基酸序列研究进展. 生物化学与生物物理进展, 1993, 20: 181～185

6 Ambler R P. Enzymatic hydrolysis with carboxypeptidases. Methods Enzymol, 1972, 143

7 Hayashi R, Moore S, Stein WH. Carboxypeptidase from yeast. Large scale preparation and the application to COOH-terminal analysis of peptides and proteins. J Biol Chem, 1973, 248(7): 2296～2302

8 Hawke D H, *et al*. Studies on C-terminal analysis. In: Hugli T C ed. Techniques in protein chemistry San Diago: Academic Press, Inc., 1989, 59～70

9 张丽华等. HPLC 在线二阶微分光谱法检测基因工程产品的 C-端肽. 科学通报, 1995, 40(14): 1314～1317

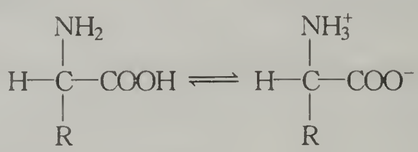
10 Perkin Elmer. Biosystems Reporter, Issue No. 24, p. 4～5, Feb. 1995. cf. Martiner, A *et al*. Biochem J, 1995, 306: 589～597

11 Boyd VL, Bozzini M, Zon G *et al* Sequencing of peptides and proteins from the carboxy terminus. Anal Biochem, 1992, 206(2): 344～352

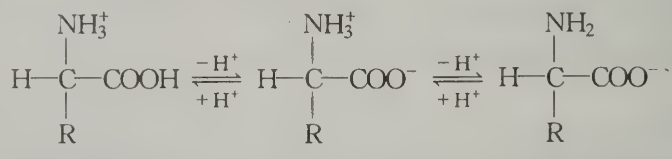
第五章　薄层等电聚焦

第一节　原　　理

蛋白质等电聚焦(isoelectric focusing,IEF)技术在20世纪60年代后才得到迅速发展和推广应用,现已成为一种被广泛采用的蛋白质分析和制备技术。与常规电泳不同之处是在IEF时,蛋白质分子是在含有载体两性电解质形成的一个连续而稳定的线性pH梯度中进行电泳。通常使用的载体两性电解质是脂肪族多氨基多羧酸(或磺酸型,或羧酸磺酸混合型),其在电泳中形成的pH范围有3～10、4～6、5～7、6～8、7～9和8～10等可应用于大多数蛋白质等电点(pI)的范围[1~3]。由于蛋白质是由不同的L－α氨基酸以不同数量和比例按一定顺序排列组成的,而氨基酸分子中的氨基和羧基都是可解离的基团,因此氨基酸可以形成两性离子:



在不同pH条件时带电荷状态会有不同:



pH 1时的主要形式　　pH 7时的主要形式　　pH 11时的主要形式

其中,R为侧链,R不同决定所带的正、负电荷和氨基酸的酸碱性质。蛋白质所带的净电荷是其组成各氨基酸残基上所有正负电荷的总和,所以蛋白质等电点是一个常数。

IEF电泳时,形成正极为酸性、负极为碱性的pH梯度。当将某种蛋白质(或多种蛋白质)样品置于负极端时,因pH＞pI,蛋白质分子带负电,电泳时向正极移动;在移动过程中,由于pH逐渐下降,蛋白质分子所带的负电荷量逐渐减少,蛋白质分子移动速度也随之变慢;当移动到pH＝pI时,蛋白质所带的净电荷为零,蛋白质即停止移动。当蛋白质样品置于阳极端时,也会得到同样的结果。因此,在进行IEF时可以将样品置于任何位置。各种不同氨基酸的蛋白质在IEF电泳结束后,会分别聚集于相应的等电点位置,形成很窄的一个区带。IEF不仅能获得不同种类蛋白质的分离和纯化效果,同时也能得到蛋白质的浓缩效果。在IEF中蛋白质区带的位置,是由电泳pH梯度的分布和蛋白质的pI所决定的,而与蛋白质分子的大小和形状无关。一般蛋白质等电点分辨率可达0.01pH单位。

第二节 薄层 IEF

聚丙烯酰胺水平薄层平板等电聚焦(horizontal thin layer polyacrylamideslafs)分析技术操作简便,样品用量少,可在一块板的相同条件下,同时分析多个样品并进行比较,电泳时间短,分辨率高,凝胶板易制成永久保存的干胶板。

制备型等电聚焦、毛细管等电聚焦将分别在有关章节中讨论。

一、仪器

实验使用 Bio-Rad 的等电聚焦仪(Model 111 Mini IEF Cell),包括制胶模具、电源(PowerPac 1000 Power Supply)(图 5.1)。



图 5.1 薄层等电聚焦仪和制胶框

二、试剂和储存液配制

①10%(m/V)过硫酸铵:100mg 过硫酸铵加水至 1ml。当天新鲜配制。

②24.25%(m/V)丙烯酰胺和 0.75%(m/V)N'-甲叉双丙烯酰胺储存液,过滤后 4℃保存,不超过一个月。

③25%(m/V)甘油:称 25g 甘油加重蒸水至 100ml。

④TEMED。

⑤两性电解质(根据需要选择 pH 范围)。

三、制胶模具的准备

取一片玻璃片(65mm×125mm)表面滴上 1~2 滴重蒸水;再取一片塑料薄膜,疏水面盖在玻璃片带水面上,使玻璃片与塑料膜贴紧并压去气泡。在凝胶制作的模具盒两侧各有 10mm 宽、0.5mm 厚的塑料条。将玻璃片上的塑料膜亲水面向下,紧靠模具盒一端两侧塑料条上贴紧,形成玻璃板在模具盒内三面贴着盒壁,只有一面敞开。

四、凝胶的制备

将以下溶液置于抽气瓶内：

重蒸水	5.5ml
丙烯酰胺/Bis 储存液	2.0ml
25%甘油	2.0ml
40%两性电解质	0.5ml
抽气 10min 后加入激活剂	
10%过硫酸铵	40μl
TEMED	10μl

轻轻混匀后用 10ml 移液管将凝胶液一次吸入管内；管口放在上述模具盒中玻璃板敞开面的边上，慢慢地将凝胶液灌注入盒内。液流需连续不断，以免形成气泡。当凝胶灌注至充满整个玻璃板框时停止。水平放置 1h，待凝胶充分聚合。聚合后玻璃板塑料膜和凝胶已连在一起。从模具盒中取出即可进行加蛋白样品和 IEF 电泳，或用保鲜膜包封闭 4℃ 冰箱保存备用。

五、加　　样

将模具中制备好的聚丙烯酰胺凝胶面朝上，凝胶表面覆盖上带有加样孔的塑料薄膜，使薄膜与凝胶贴紧，每孔蛋白质样品（溶于 50%甘油中）可以加入 0.5～2μl，样品量超过 2μl 可将样品滴在滤纸。加样后需 5～10min，待样品进入胶后小心揭去带加样孔的膜（如滤纸可在 IEF 结束后再揭去）。

六、电　　泳

将电泳盒内的石墨电极清洗干净，用水将电极润湿，将加有样品的凝胶面向下，直接将凝胶的两端放在石墨电极上，盖好电泳槽盖，即可开始电泳。

在稳压条件下，开始聚焦在 100V，15min 后增加电压至 200V；聚焦 15min 后再增加电压至 450V；60min 便可以结束电泳。

七、染色和脱色

①固定液：12.5%三氯乙酸，30%乙醇。

②染色液：0.04%考马斯亮蓝 R250（Coomassic brilliant blue R250）；10%冰醋酸；27%乙醇。

③脱色液：25%乙醇；7%冰醋酸。

IEF 结束后可用微量表面电极测定 p*I*。如凝胶需染色，可将凝胶连同塑料薄膜一起从玻璃板上取下。

八、凝胶保存

脱色完毕后，轻轻地从玻璃上取下支持膜，室温下干燥。此胶可永久保存(图5.2)。

第三节　等电聚焦中的注意事项

①pH梯度的选择：可先在宽pH范围内(pH3.5～10或pH2.5～11)的载体两性电解质中进行分析IEF。当了解到目的蛋白质的pI值后，再用窄pH范围(pH3～5或pH5～7)的载体两性电解质进行分析或制备IEF。在窄pH范围进行IEF不仅可以提高分辨率，而且可以增加需分离蛋白质的负载容量。

②IEF如在宽pH范围载体两性电解质内进行，为了克服在中性区域形成纯水区带，可适当添加中性载体两性电解质。

③为防止电泳过程中pH梯度的衰变，一般在电流降低达最小而恒定时尽快结束IEF。

④pH梯度(pI)测定。

凝胶IEF后，可分段切割凝胶，用3～5倍体积的蒸馏水或10mmol/L KCl浸泡该凝胶，从凝胶浸出液中测定pH值；或用微电极直接测定凝胶表面pH值。

薄层等电聚焦后，可用微电极检测凝胶表面pH值；根据蛋白质分布的pH值即能确定该蛋白质的pI。

⑤IEF聚焦过程中由于蛋白质泳动到某一区段，而该区段刚好是该蛋白质的pI，会造成蛋白质的沉淀。为了解决此问题，可在样品中添加脲、TritonX-100、NP-40或其他一些非离子型表面活性剂。

⑥IEF样品应不含盐类。因盐浓度高在高压电泳时电流大，发热量大；但在无盐时，一些蛋白质的溶解度低，可通过在样品中加入适量两性电解质或甘氨酸来解决。

⑦所用试剂要求高纯度。

第四节　IEF的优缺点

(1) 与常规电泳比较IEF有以下优点

①具有很高的分辨率，可把pI相差0.01pH单位的蛋白组分相互分开。

②样品体积和加样位置不受严格限制，样品可混入胶中或放在任何位置。

③在IEF电泳一结束，即可以测定其蛋白质pI。

④分离蛋白质快速，一般分析型薄层IEF可在2h完成。

⑤由于载体两性电解质分子质量小，不会与蛋白质反应并使之变性，因此只需硫酸铵沉淀，通过透析和分子筛、电泳等方法很易与蛋白质分开。

⑥IEF分离获得的蛋白质基本上可保持原有的生物活性。

(2) IEF的缺点：

①许多蛋白质在pI附近会产生沉淀，因而影响分离效果。可在凝胶介质中或制备

电泳无介质的样品溶液中加脲或非离子去垢剂解决。

②样品中最好不含盐。如高盐浓度在高压电泳时,电流大,发热量大;但在无盐时,一些蛋白质溶解不好;这可通过在样品中适当多加些两性电解质或甘氨酸等两性物质来解决。

③由于两性电解质可与蛋白质间的静电作用而紧密结合,形成的复合物可能导致它与抗体反应降低或使酶活性降低。可通过高离子浓度(0.1mol/L 或更高,达 0.5mol/L)、pH7 左右的挥发性缓冲液破坏它们的静电作用。

表 5.1 提供了部分蛋白质的等电点和分子质量。

表 5.1　部分蛋白质的等电点和分子质量

蛋白质名称	分子质量/Da	等电点(p*I*)	蛋白颜色
碳酸酐酶	38 000	4.5	
人胎盘碱性磷酸化酶	116 000	4.6	
藻青蛋白	232 000	4.63	蓝色
β-乳球蛋白	18 400	5.1	
大肠杆菌 L-天冬酰胺合成酶	80 000	5.5	
人血清白蛋白	69 000	5.85	
牛碳酸酐酶	31 000	6.0	
大肠杆菌 L-高丝氨酸脱氢酶	360 000	6.1	
人碳酸酐酶	28 000	6.5	
马肌红蛋白(二条带)	175 000	6.8,7.0	棕色
人血红蛋白 A	64 500	7.1	红色
猪红细胞碳酸酐酶	30 375	7.3	
人血红蛋白 C	64 500	7.5	红色
扁豆凝集素(三条带)	49 000	7.8,8.0,8.2	
豚鼠胰 α 淀粉酶	52 000	8.4	
细胞色素 c	12 200	9.6	红色
鼠肝醇脱氢酶	68 000	9.7	
菠萝蛋白酶	22 000	9.7	

第五节　实　　例

①用 IEF 来鉴定蛋白质不同纯化过程所得到的样品(图 5.2)。



图 5.2　不同蛋白质样品的薄层等电聚焦

用此法可以来鉴定不同纯化方法所得到的蛋白质的性能，根据需要以此来选择最佳的纯化方法。

②用 Rotofor 分离纯化样品(可参见第七章第三节)，用 Mini-IEF 来鉴定(图 5.3)。

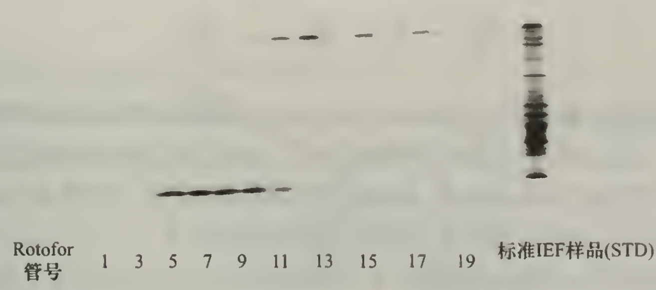


图 5.3　细胞色素 c 和白蛋白的制备

IEF 分离(Bio-Rad Rotofor)，然后用 Mini-IEF 鉴定

自由流等电聚焦后样品，也可通过薄层等电聚焦来鉴定分离纯化效果。

(邵晓霞　龚蓁蓁)

参考文献

1　Vesterberg O, Svensson H. Isoelectric fractionation, analysis, and characterization of ampholytes in natural pH gradients. IV. Further studies on the resolving power in connection with separation of myoglobins. Acta Chem Scand. 1966, 20(3): 820～834

2　Vesterberg O. Buffering capacity and conductance of isoionic ampholytes. Acta Chem Scand, 1969, 23: 2653

3　Vesterberg O. Physicochemical properties of the carrier ampholytes and some biochemical applications. Acta Chem Scand, 1973, 27: 2415～2420

第六章　SDS－聚丙烯酰胺凝胶电泳

第一节　概　　况

聚丙烯酰胺凝胶电泳(polyacrylamide gel electrophoresis,PAGE)具有较高的分辨率和灵活性,因而被广泛应用于蛋白质的分离和分析。然而,凝胶电泳目前仍是一项手工操作较多的技术,操作者如能对电泳原理、聚胶过程及影响电泳的一些因素多加了解,将有助于顺利完成有关实验,获得分离效果及重复性好的电泳结果。

聚丙烯酰胺凝胶是由丙烯酰胺和甲叉双丙烯酰胺(bis)经共聚合形成。此聚合过程是由四甲基乙二胺(tetramethylethylenediamine,TEMED)和过硫酸铵激发的,即过硫酸铵在水溶液中形成一个过硫酸自由基,然后此自由基再激活 TEMED,随后 TEMED 作为一个电子载体提供一个未配对电子将丙烯酰胺单体转化成一个自由基从而激活了后者。被激活的单体和未被激活的单体反应开始了多聚链的延伸,正在延伸的多聚链也可以随机地通过 bis 的作用进行交叉互联成为网状结构,最终多聚物聚合成凝胶状,而其孔径大小等特征由聚合条件及单体浓度决定。凝胶的孔径可以在一较宽范围内变化以迎合不同的分离需要,改变凝胶或缓冲液的某些组成成分,就可以按照不同的分离机制进行分离,如分别根据蛋白质的电荷、大小或荷质比特性进行等电聚焦、SDS-PAGE 及酸性或碱性凝胶电泳等。

核黄素也可以作为产生自由基的起始物,有时也将它和过硫酸铵混合在一起使用,核黄素激活聚合反应需要光和氧气存在,因此称为光化学聚合。

在聚丙烯酰胺凝胶电泳中最常用的是 SDS-PAGE(十二烷基硫酸钠－PAGE),简称 SDS 电泳。SDS 是一种阴离子去污剂,能断裂分子内和分子间的氢键,它与蛋白质的疏水部分相结合,破坏其折叠结构,并使其稳定地存在于均一的溶液中。SDS 蛋白质络合物的长度与其分子质量成正比。由于在样品介质和聚丙烯酰胺凝胶中加入了离子去污剂和强还原剂(如巯基乙醇和二硫苏糖醇),使 SDS 蛋白质络合物所带的负电荷大大超过了蛋白质分子原有的电荷量,消除了不同分子之间原有的电荷差别,其电泳迁移率主要取决于亚基分子质量的大小。当蛋白质的分子质量为 15 000～200 000Da 时,电泳迁移率与分子质量的对数呈线性关系。因此,它被广泛应用于蛋白质纯度鉴定和分子质量测定,且因所需设备简单,操作简便,重复性好,使它成为许多研究领域中一种重要的分析和微量制备技术。

第二节　材　　料

一、设备和材料

①电泳装置为 Bio-Rad Mini-Protean II 型电泳仪。

②电源(Bio-Rad Model 3000xi)。
③水浴锅。
④Eppendorf 管。
⑤移液器(1000μL、200μL，20μL)，Tip。
⑥摇床。
⑦水泵。
⑧干胶仪(Bio-Rad Model 583 Gel Dryer)。
⑨玻璃纸，Whatman 3MM 滤纸。
⑩pH 计。
⑪离心机。

二、试　　剂

①丙烯酰胺(电泳级)。
②*N*，*N*′-甲叉双丙烯酰胺(bis)。
③SDS。
④TEMED。
⑤Tris。
⑥过硫酸铵。
⑦α-巯基乙醇。
⑧甘油。
⑨甘氨酸。
⑩溴酚蓝。
⑪盐酸。

第三节　基本操作步骤

一、聚胶前准备

①配制储存液，单体储存液需经 Millipore 膜过滤后使用。
②配制 10%过硫酸铵溶液(需每天新鲜配制)。
③将单体储存液、缓冲液、水按照一定比例配制成凝胶溶液。
④将灌胶装置准备好。
⑤待凝胶溶液平衡至室温(23～25℃)后，真空脱气 15min。

二、不连续系统下层分离胶的聚合

在 10ml 凝胶溶液中加入 10%过硫酸铵 50μl(最终浓度为 0.05%)，TEMED 原液 5μl(最终浓度为 0.05%)。轻缓地摇动溶液(8～10 下)，使激活剂混合均匀。

将凝胶溶液平稳地注入两层玻璃板中(注意要以连续平稳的液流从中间位置注入),再在液面上小心注入一层水或正丁醇(2～3mm 高),以阻止氧气进入凝胶溶液中,然后静置至少 90min。

三、不连续系统中的上层浓缩胶和连续系统凝胶的聚合

连续系统只含一种凝胶,它和不连续系统中的浓缩胶具有相同之处,即在其液面处与空气相接触,因此激活剂的含量应相应地增加,在 10ml 凝胶溶液中加入 10%过硫酸铵 50μl(最终浓度 0.05%),TEMED 原液 10μl(最终浓度 0.1%)。

同前将激活剂混合均匀,但摇动溶液时不要摇得过于剧烈以免引入过多的氧气。

吸去不连续系统中下层分离胶上的水分,以连续平稳的液流注入凝胶溶液,然后小心插入梳子并注意不得在齿尖留有气泡。

静置 90min 以上以保证完全聚合。

如果在聚胶过程中发现一道道纹线,即凝胶不均匀,可能是聚合速度太快或激活剂未充分混合导致局部浓度过高。

凝胶聚合的速度可以通过紫外吸收检测,由于丙烯酰胺中的双键在 260nm 有吸收,聚合后双键消失,光吸收减低。将凝胶溶液倒入石英比色杯中,可观察到 20～30min 后光吸收读数开始下降,至 80min 以后趋于稳定,说明聚合完成。

四、加　　样

将聚合好的凝胶板安置于电泳槽中,加入上下槽电泳缓冲液后,小心拔去梳子,以缓冲液冲洗加样槽后,依次加入标准品和待分析样品。注意加样时间要尽量短,以免样品扩散。

五、电　　泳

加样完毕,盖好电泳槽盖子并选择适当的电压进行电泳。通常,在不连续系统中,上层浓缩胶的电泳电压要低于分离胶的电泳电压,使样品更好地进入凝胶。当指示剂即将走出凝胶说明电泳基本结束。

六、染色和脱色

电泳完毕需通过染色和脱色评定其结果的优劣。此外,根据不同的研究目的,也可将凝胶直接进行放射自显影及电印迹。以下介绍最常用的染色和脱色方法。

1. 考马斯亮蓝染色

染色液为 0.1%考马斯亮蓝 R-250,40%甲醇(或乙醇)和 10%冰醋酸,染色 0.5～lh。

脱色液为 10%甲醇(或乙醇)和 10%冰醋酸,需 1～3h,其间更换 2～3 次脱色液至背

景清楚。

此方法检测灵敏度为 0.1～1.0μg[1]。

2. 银染法

根据 Morrisey[2]方法并稍加改进，具体步骤如下：

第一步：电泳结束后凝胶在 50% 甲醇和 10% 冰醋酸中浸泡 30min，然后换至 5% 甲醇和 7% 冰醋酸中 30min。

第二步：10% 戊二醛固定 30min。

第三步：重蒸水中漂洗过夜或 2h（中间更换三次水），这一步根据电泳结束的时间选择。

第四步：在 5μg/ml 硫苏糖醇（DTT）溶液中浸泡 30min。

第五步：在 0.1% $AgNO_3$（定影液）中浸泡 30min。

第六步：在水中迅速漂洗几秒钟。

第七步：用显影液（100ml 3% Na_2CO_3 中含 50μl 37% 甲醛，经 Millipore 膜过滤）漂洗两次。

第八步：在显影液中浸泡至蛋白条带显现。

第九步：在 50ml 显影液中加入 5ml 2.3mol/L 的柠檬酸终止显影过程。凝胶需在此终止液中至少浸泡 10min。

第十步：重蒸水漂洗三遍，共 30min。

第十一步：凝胶在 0.03% Na_2CO_3 溶液中浸泡 10min，使凝胶处于微碱性环境中，蛋白质条带不易褪色。

第十二步：凝胶在 3% 甘油中浸泡 30min 以增加柔软性，然后用干胶机或简单地夹在 2 张事先润湿的透气玻璃纸中（中间不得留有气泡），用固定夹封住制成凝胶干片。

需注意以上每一步凝胶在溶液中都要在水平摇床上轻轻摇动。



图 6.1 G-CSF 纯度鉴定（银染）

1,2,3 为不同批号的重组 G-CSF 用巯基乙醇处理后的电泳条带；6,7,8 为不同批号的重组 G-CSF 未经巯基乙醇处理的电泳条带；4,5 为标准蛋白质

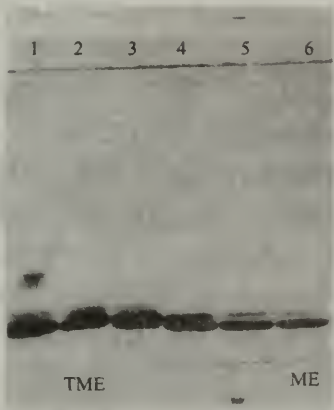


图 6.2 GM-CSF 纯度鉴定（银染）

1,2,3 为不同浓度 GM-CSF 用巯基乙醇处理后的电泳条带；4,5,6 为不同浓度 GM-CSF 未经巯基乙醇处理的电泳条带

此方法检测灵敏度高于考马斯亮蓝法，为 2～10ng[3]，常可用于鉴定蛋白质的纯度（见图 6.1，图 6.2），但是由于不同蛋白质结合银离子的能力不同，而且在上述第八步中显影时间的长短将明显改变蛋白条带的色度，所以，应用银染法时根据蛋白质条带深浅确定其浓度不易准确。

3. 铜、锌等金属离子染色法

电泳结束后凝胶在水中漂洗几秒钟，然后浸泡在 0.3mol/L 的 $CuCl_2$（或 $ZnCl_2$ 等）溶液中，并在室温下轻摇 5min，这时若在凝胶背面衬以暗色背景，即可观察到近乎透明的蛋白质条带的显现，用水冲洗 2～3min 除去过量的染色剂后，可保存在重蒸水中[4,5]。

七、干　　胶

将一张经润湿的玻璃纸（或 Whatman 3MM 滤纸）平整地铺于干胶仪桌面上，凝胶置于玻璃纸上，以蒸馏水作润滑剂，再将另一张同样经润湿的玻璃纸覆盖在凝胶的表面，小心去除所有的气泡，盖上密封垫膜，打开真空开关使密封垫膜的四周密封于干胶仪上，加盖后开启电源进行程序干胶。

凝胶也可在室温和常压下，置于玻璃纸内，四周以封条密封（不能留有气泡以免破裂）后进行自然干胶。

第四节　几种不同凝胶电泳系统及其应用

在聚丙烯酰胺凝胶电泳中最常用的是 SDS-PAGE，它被广泛应用于蛋白质纯度鉴定和分子质量测定，下面介绍适用于不同范围蛋白质分离的实验方法。

一、Laemmli 系统

贮存液由以下几部分组成。

①贮备胶（30%T，2.67%C）：称取 29.2g 丙烯酰胺和 0.8g bis，用水溶解并定容至 100ml，Millipore 膜过滤后在 4℃ 避光保存。

②1.5mol/L Tris-HCl，pH 8.8：称取 18.15g Tris，用 60ml 重蒸水溶解，再用 1mol/L 盐酸调节至 pH8.8，然后定容至 100ml，4℃ 保存。

③0.5mol/L Tris-HCl，pH 6.8：称取 6g Tris，用 60ml 重蒸水溶解，再用 1mol/L 的盐酸调节至 pH6.8，然后定容至 100ml，4℃ 保存。

④10%SDS：将 10g SDS 溶于 80ml 重蒸水中并轻缓搅拌（室温低时需水浴加热溶解），再定容至 100ml，室温存放。

⑤样品缓冲液：分别取重蒸水 4.0ml，0.5mol/L Tris-HCl，pH6.8 缓冲液 1.0ml，甘油 0.80ml，10%（m/V）SDS 1.6ml，巯基乙醇 0.4ml，0.1%（m/V）溴酚蓝 0.2ml，总体积为 8ml，加样前样品与此缓冲液以 1∶4 混合后在沸水浴中加热 4min。

⑥浓缩电泳缓冲液（5×）：称取 15g Tris，72g 甘氨酸，5g SDS，用水溶解至 1 000ml，

pH8.3，不必调节。用时稀释5倍。

电泳前首先根据样品分子质量大小选择适宜的凝胶浓度，当分析一个未知样品时，常先用7.5%的标准凝胶或用4%～10%的凝胶梯度来测试，然后选出适宜的浓度。

蛋白质分子质量范围/10^3Da	适宜的凝胶浓度/%T
<10	15～20
10～50	15
15～60	12.5
15～100	10
30～120	7.5
30～200	5
>200	2～5

分离胶配方

	15%T	12%T	10%T	7.5%T
重蒸水	2.35ml	3.35ml	4.02ml	4.85ml
1.5mol/L Tris-HCl，pH 8.8	2.50ml	2.50ml	2.50ml	2.50ml
30%丙烯酰胺/bis	5.0ml	4.0ml	3.33ml	2.50ml
	室温脱气15min			
10%SDS	100μl	100μl	100μl	100μl
10%过硫酸铵(新鲜配制)	50μl	50μl	50μl	50μl(0.05%)
TEMED	5μl	5μl	5μl	5μl(0.05%)
总体积	10ml	10ml	10ml	10ml

浓缩胶配方(4.0%T，0.125mol/L Tris，pH6.8)

重蒸水	3.05ml
0.5mol/L Tris-HCl，pH6.8	1.25ml
30%丙烯酰胺	0.65ml
室温脱气15min	
10%SDS	50μl
10%过硫酸铵	25μl(0.05%)
TEMED	5μl (0.1%)
总体积	5ml

在Laemmli系统中蛋白质分子质量的对数值和其相对迁移距离呈良好的线性关系，因此适合测定蛋白质的分子质量。根据一系列已知分子质量蛋白质的相对迁移距离制得标准曲线，将未知蛋白质的相对迁移值对照标准曲线即可获得其分子质量数据(图6.3)。

另外，将标准曲线向Y轴(log分子质量)延伸可得Y轴截距，即相对泳动为零时的分子质量值，也就是说当蛋白质的分子质量大于此值时将不能进入分离胶中，Y轴截距的数



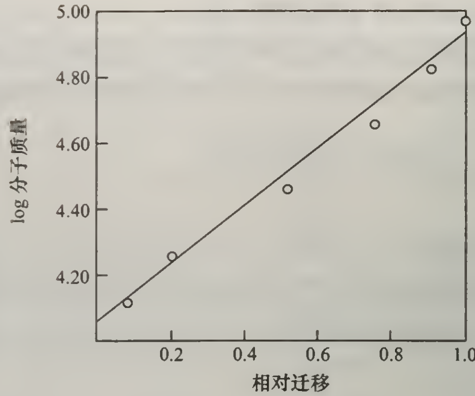


图 6.3　标准蛋白质及重组 IL－3 产品鉴定(考马斯亮蓝染色)

1～5 为不同批号的重组白细胞介素－3;6 为标准蛋白质

值大小可体现聚胶完全程度以及电泳结果的重复性。

表 6.1 列出了商品化的标准蛋白质及它们的分子质量。

表 6.1　标准蛋白质及其分子质量

蛋白质	分子质量 /10^3Da	蛋白质	分子质量 /10^3Da
胰岛素 A 链、B 链(insulin A chain & B chain)	2,3	内切糖苷酶 H(endoglycosidase H)	29
牛胰蛋白酶抑制剂标准蛋白质及其分子质量(bovine pancreatic trypsin inhibitor)	6.2	乳糖脱氢酶(lactate dehydrogenase)	36
HIV 蛋白酶(HIV protease)	10	3－磷酸甘油醛脱氢酶(glyceraldehyde-3-phosphate dehydrogenase)	36
细胞色素 c(cytochrome C)	13	卵白蛋白(ovalbumin)	45
α－乳白蛋白(α-lactalbumin)	14.2	γ－球蛋白重链(γ-globulin heavy chain)	50
溶菌酶(Lysozyme)	14.3	过氧化氢酶(catalase)	57
核糖核酸酶 A(ribonuclease A)	14	牛血清白蛋白(bovine serum albumin)	66
抹香鲸肌红蛋白(sperm whale myoglobin)	17.2	转铁蛋白(transferrin)	75
β－乳球蛋白(β-lactoglobulin)	18.4	伴白蛋白(conalbumin)	77
大豆胰蛋白酶抑制剂(soybean trypsin lnhibitor)	20.1	磷酸化酶 b(phosphorylase b)	97.4
γ－球蛋白轻链(γ-globulin light chain)	25	β－半乳糖苷酶(β-galactosidase)	116
胰蛋白酶原(trypsinogen)	24	α2－巨球蛋白(α2-macroglobulin)	180
α－胰凝乳蛋白酶原(α-chymotrypsinogen)	25.5	肌球蛋白(myosin)	205
碳酸酐酶(carbonic anhydrase)	29		

二、大孔胶系统

在上述 Laemmli 系统中,交联度 C% 为 3%(3% 为 bis 的质量/单体的总质量),但是有时样品中的蛋白质较复杂时,在一块单一浓度凝胶上很难将所有蛋白分离,比如单体浓

度 T 为 12% 的凝胶就不能分离 100kDa 以上的蛋白质。虽然梯度凝胶可以拓宽分子质量分离范围,但操作繁琐。以下介绍一种大孔径凝胶,它可以在一块凝胶上分离分子质量 $10\times10^3\sim450\times10^3$Da 的蛋白质,比如细胞内总蛋白分析。这种凝胶还可以允许上样量高达 200μl,并仍能得到尖窄的条带。另外,这种凝胶中的蛋白质更容易转移至膜上。但是由于大孔胶系统中蛋白质分子质量的对数值和其相对迁移距离的关系的线性程度不佳,用它测定蛋白质分子质量的准确度不及常规的 Laemmli 系统。

大孔胶的配方:

分离胶中含 10% 丙烯酰胺, 0.1% bis, 0.4% SDS, 5% 甘油, 200mmol/L Tris, 100mmol/L 甘氨酸(不必调节 pH),最终为 pH9.0。聚胶时加入 0.1%过硫酸铵和 0.05% TEMED。

浓缩胶中含 4% 丙烯酰胺, 0.04% bis, 0. 4% SDS, 4mmol/LEDTA, 5% 甘油, 70mmol/L Tris-HCl, pH 6.7。聚胶时加入 0.1%过硫酸铵和 0.05%TEMED。

上槽(负极)电泳缓冲液为 100mmol/L Tris, 150mmol/L 甘氨酸, 0.1%SDS, pH8.45(不必调 pH)。下槽(正极)电泳缓冲液为上槽电泳缓冲液稀释一倍。电泳时选择恒电压 60V。

样品缓冲液中含 8mol/L 尿素, 3%SDS, 100mmol/L DTT, 0.005%溴酚蓝, 70mmol/L Tris-HCl, pH6.7。上样前于沸水浴中加热 4min。

三、Tris-Tricine 系统

此系统适合小分子蛋白和多肽分离,分子质量可低至 3 000Da(图 6.4)。

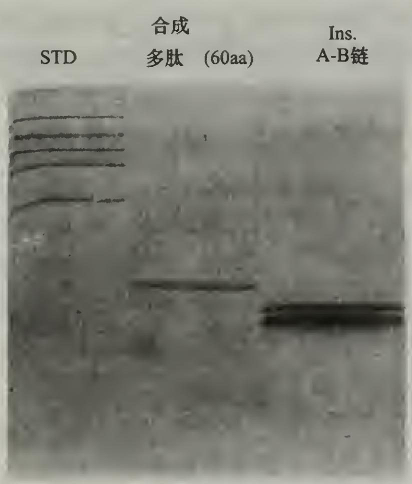


图 6.4 小分子蛋白质和多肽的分离
(用 Tris-Tricine 系统)

电泳前准备下列溶液:

①正极(下槽)缓冲液(10×, 2.0mol/L Tris-HCl, pH8.9):将 121.9g Tris 溶于 400ml 重蒸水中,用 1mol/L HCl 调至 pH8.9,再定容至 500ml。

②负极(上槽)缓冲液(10×,1.0mol/L Tris,1.0mol/L Tricine,1.0%SDS,pH8.25):将60.55g Tris,89.58g Tricine及5g SDS溶于400ml重蒸水中,用1mol/L HCl滴定至pH8.45,加水至最终体积为500ml。

③丙烯酰胺溶液(49.5%T,3%C):将48g丙烯酰胺和1.5g bis溶于100ml重蒸水中。

④凝胶缓冲液(3×):将181.5g Tris,1.5g SDS溶于400ml重蒸水中,用1mol/L HCl滴定至pH 8.45,再加水稀释至500ml。

⑤50%(*m/V*)甘油:称取250g甘油加重蒸水混合均匀并定容至500ml。

以上溶液需经Millipore膜过滤。

⑥样品缓冲液:将1ml甘油,0.5ml巯基乙醇、0.4g SDS,1ml 0.5mol/L Tris-HCl,pH6.8及少许溴酚蓝(至溶液呈棕红色)混合后再加入重蒸水定容至5ml。上样前将样品与相同体积的样品缓冲液混合后,沸水浴加热2~4min。

凝胶的配方

	10%分离胶	4%浓缩胶
单体溶液(49.5%T,3%C)	3.0ml	0.5ml
50%甘油	4.0ml	
凝胶缓冲液(3×)	3.0ml	2.5ml
重蒸水	3.92ml	4.46ml
	脱气15min	
10%过硫酸铵	75μl	37.5μl
TEMED	7.5μl	3.8μl
总体积	14ml	7.5ml

Tris-Trcine系统电泳条件根据不同仪器可在100~200V之间选择。

第五节 注意事项

一、形成凝胶的试剂要有足够的纯度

丙烯酰胺是形成凝胶溶液中的最主要成分,其纯度的好坏将直接影响凝胶的质量,丙烯酰胺可能混有的影响凝胶形成的杂质有:丙烯酸、线性多聚丙烯酰胺、金属离子等。

丙烯酸是丙烯酰胺的脱氨产物,它也能和丙烯酰胺及bis发生共聚合,从而导致凝胶中部分区域的pH发生改变,最终电泳结果产生偏差,如表现为:迁移位置发生改变、蛋白条带弯曲不平整等。除了固体丙烯酰胺中可能含有丙烯酸外,单体储存时间过长(4℃,一个月以上),也将有部分丙烯酰胺水解成丙烯酸。

线性聚丙烯酰胺是丙烯酰胺在生产、加工或储存时污染了具有催化作用的杂质引起"自聚合",凝胶溶液中混有这种丙烯酰胺的聚合物后将造成聚胶不均匀并改变凝胶的孔径,从而影响电泳的重复性。

金属离子的含量超过一定浓度后有时会抑制凝胶的聚合,有时则改变一些蛋白质(尤其是金属结合蛋白)的泳动力。

上述杂质虽然可以通过一些物理、化学方法检测，但通常都是根据聚胶及电泳结果推测的。在实验室中可以通过重结晶的方法纯化丙烯酰胺，但由于丙烯酰胺单体具有神经毒性，一般应尽量避免上述工作。最简单的方法是选择质量好、达到电泳纯级的产品。另外，需注意不能随便倾倒单体溶液，应加入过量的催化剂使其完全聚合成无毒性的聚丙烯酰胺后再弃置。

甲叉双丙烯酰胺相对于丙烯酰胺在单体溶液中含量甚少，但也可能混有上述几种杂质，造成同样的结果。

TEMED 中混有氧化试剂后其自身极易被氧化而失去催化能力，另外 TEMED 的氧化产物呈黄色，以此可以鉴定 TEMED 的质量。如果无法获得无色透明的 TEMED，只能适当增加用量以保证聚胶速度。

过硫酸铵容易吸潮，由于它溶解于水后很快降解失去催化能力，因此潮解后的过硫酸铵会渐渐失去活性。保存固体过硫酸铵应密闭干燥。过硫酸铵溶液宜制胶当天配制。另外在配制凝胶溶液时过硫酸铵不易过量，否则会使蛋白质在电泳过程中被氧化（主要指天然无变性剂凝胶）。

凝胶溶液中的其他成分如 Tris、甘氨酸、SDS、尿素等也应选择纯度较高的产品（分析纯以上），以避免混有抑制凝胶聚合的杂质。

二、激活剂的用量

激活剂的浓度适当与否不仅可以决定凝胶聚合的速度，也将影响凝胶的质量。增加过硫酸铵或 TEMED 的用量，将使丙烯酰胺聚合链长度缩短，凝胶会变得脆弱而失去应有的柔性。二者多至一定程度时，聚合链长度过短，将不会形成肉眼可见的凝胶，这些短链多聚物呈溶液状，只是其黏度较聚合前高一些。过量的过硫酸铵或核黄素还将氧化蛋白质尤其是含有巯基的蛋白质，使它们的泳动力发生改变；而过量的 TEMED 不但可以与蛋白质反应使其电泳条带位置偏移，还将升高缓冲液的 pH。

激活剂过少则聚胶速度过慢，氧气越来越多地进入单体溶液中抑制聚合过程，最终凝胶孔径扩大，支撑力弱而易于变形。另外，未聚合的丙烯酰胺和 bis 单体还可以和蛋白质中的 α-氨基、巯基、酚羟基等反应，造成蛋白质泳动力改变，条带拖尾或停滞在加样孔。

由此可见，可以通过激活剂的用量控制聚胶速度，保证聚胶完全并提供理想的凝胶质地。

对于非连续系统中的下层分离胶最佳聚合速度是从加入过硫酸铵和 TEMED 起至开始出现凝胶的时间为 15～20min（并不是此时已凝聚完全）。对于浓缩胶及不含浓缩胶的连续系统，最佳聚合速度为 8～10min 开始可见聚合。这种情况凝胶与大气接触，需要增加激活剂的浓度并加快聚合速度。总之，要掌握一个原则，即用尽量少的催化剂并在最佳时间聚合。

使用核黄素的优点是只需要很少量（5～10μg/L）就有催化活性，尤其适合于缓冲能力较弱的凝胶系统，如只含有两性电解质的等电聚焦凝胶。另外，核黄素具有抗氧化能力，通常用它作为激活剂时，聚合物在 30～60min 后开始呈凝胶。

三、温度的影响

聚胶的最佳温度为23～25℃，需要注意灌胶前单体溶液(通常置于4℃冰箱)及玻璃板或管(有时从烘箱取出)也要平衡至此温度。由于溶液在抽真空时仍维持较低温度，需待其升至室温后再脱气。另外，升温后脱去氧气的速度较低温时快。

四、氧气的影响

氧气会与被激活的单体自由基作用抑制聚合过程，因此需在加激活剂前对单体溶液脱气。如果省去这一步骤，凝胶仍能聚合但需更多的激活剂，并且不能保证很好的重复性，充分去除氧气可以用尽量少的激活剂得到最佳聚合效果。通常在较好的真空度($\leqslant$125torr，1torr＝1mmHg＝1.33322×10^2Pa)脱气至少15min。

使用核黄素作为激活剂时(核黄素/TEMED系统或核黄素/TEMED/过硫酸铵)，由于核黄素由无催化活性状态转化成具有催化作用的形式时需要少量氧气，脱气不能太完全，一般不得超过5min。

五、凝胶添加剂

凝胶中常常加入SDS、尿素、Triton X-100等添加剂。SDS加入后不会对凝胶的聚合产生影响，但尿素会使凝胶孔径变小，这一特性可用来分离小分子蛋白质或多肽。

六、聚胶时间

虽然应用过硫酸铵/TEMED催化后灌胶15～20min即可观察到凝胶的形成，但至少需90min才能保证95%以上的单链聚合成长链以及具有合适的孔径大小。采用核黄素时聚胶时间更长，但它一般用于等电聚焦即根据蛋白质的电荷性质进行分离，对孔径大小要求不高，因此不必等很长时间(完全聚合需8h)。

七、单体的浓度%T，%C

所谓单体浓度%T是指单体总质量(丙烯酰胺＋双丙烯酰胺)在溶液中的百分比(m/V)，可选择的范围为3%～30%。单体浓度增加后聚合速度将加快，因此可适当减少催化剂的用量。

%C是指双丙烯酰胺与单体总质量的比值(m/m)。

八、操作注意点

①未聚合的丙烯酰胺具神经毒性，操作时应戴手套防护。

②梳子插入浓缩胶时,应确保没有气泡,可将梳子稍微倾斜插入以减少气泡的产生,梳子拔出时亦应小心,不要破坏加样孔。

③电泳槽内加入电泳缓冲液后,清除黏附在凝胶底部的气泡,以保证电泳过程中电流的畅通。

④加样前可先用电泳缓冲液冲洗加样孔,这样可去除未聚合的丙烯酰胺。

⑤加样前样品应先离心,尤其是长时间放置的样品,以减小蛋白质条带的拖尾现象。

⑥为避免边缘效应,可在未加样的孔中加入等量的样品缓冲液。

⑦样品缓冲液中煮沸的样品可在-20℃保存几周,但反复冻融会使蛋白质降解。

⑧上样时,小心不要使样品溢出而污染相邻的加样孔。

⑨电泳时,应采用恒压的模式,这样蛋白质才可以保持恒定的电泳迁移率[9]。低电压可使电泳过程中产热减小,迁移减慢,同时分辨率会稍有下降。

⑩为减小蛋白质条带的扩散,上样后应尽快进行电泳,电泳结束后也应直接进行固定或染色。

⑪考马斯亮蓝染色液应过滤后使用,以免未溶解的染料沉积于胶上。

⑫干胶前可将凝胶浸泡于5%的甘油中过夜,这样可减小干胶过程中出现胶破裂[10]。

（宋金芳　屠红旻）

参考文献

1 Smith I, Cromie R, Stainsby K. Seeing Gel Wells Well. Anal Biochem, 1988, 169:370～371

2 Morrissey, JH. Silver stain for protein in polyacrylamide gels: a modified procedure with enhanced uniform sensitivity. Anal Biochem, 1981, 117:307～310

3 Giulian GG, Moss RL, Greaser M. Improved methodology for analysis and quantition of proteins on one-dimensional silver-stained slab gels. Anal Biochem, 1983, 129:277～287

4 Dzaudu JF, Johnson JF, Wise GE. Sodium dodecyl sulfate gel electrophoresis: staining of polypeptides using heavy metal salts. Anal Biochem, 1988, 174:157～167

5 Lee C. Copper staining: a five-minute protein stain for sodium dodecyl sulfate polyacrylamide gels. Anal Biochem, 1987, 166;308～312

6 Laemmli UK. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature, 1970, 227:680～683

7 Doucet JP, Trifaro JM. A discontinuous and highly porous sodium dodecyl sulfate polyacrylamide slab gel system of high resolution. Anal Biochem, 1988, 168:265～271

8 Schagger H, Van Jagoe G. Tricine-sodium dodecyl sulfate polyacrylamide gel electrophoresis for separation of proteins in the range from 1 to 100 kDa. Anal Biochem, 1987, 166:368～373

9 Hames BD, Rickwood D. Gel electrophoresis of proteins. In: A practical approach. Oxford and Washington D.C: IRL Press. 1981

10 Walker JM. Methods in molecular biology. Volume: Proteins. New Jersey: Humana Press, 1984

第七章　液相电泳

电泳被定义为荷电物质(离子)在电场中因受吸引或排斥而引起的差速运动。1925年瑞典 Tiselius 继承前人在电泳方面的工作,在 30 年代成功地将一个无色、移动的蛋白质界面的折光率变化定量地记录下来,从而把液相界面移动电泳建立成一个分析方法,并于 1948 年获得了诺贝尔化学奖。随后,许多科学家在此基础上发展了有支撑介质的区带电泳,例如纸电泳、琼脂电泳、淀粉凝胶电泳等,直到现在应用最广泛的 SDS-PAGE。区带电泳的优点是设备简单、价廉,分析量在微克水平上,但也牺牲了液相界面电泳的一些优点。80 年代后科学家们又考虑到液相电泳的若干特点,相继发展了超微量、高分辨率的毛细管电泳和制备级的连续自由流电泳以及液相等电聚焦。我们实验室近年来相继建立和发展了这几种电泳技术,并统称之为"第二代液相电泳"[1],下面分节逐一介绍。

第一节　毛细管电泳

一、概　　况

1. 发展史

在像 Tiselius 所用的自由溶液中的电泳方式下,分离效率受热扩散和对流的限制。因此,随后发展的各种电泳都在抗对流介质(如纸电泳、凝胶电泳)中进行。尽管凝胶电泳是应用最广泛的分离技术之一,但分析时间长、效率低及检测和自动化困难等依然是自身的缺陷。由于细径毛细管自身具有抗对流作用,可以替代原来起此作用的凝胶介质,因此近年来细孔管或毛细管被用于电泳。20 世纪 80 年代初,很多科学家致力于发展微量的、没有支撑介质的液相毛细管电泳[2~5],并揭示了毛细管电泳作为一种分析技术所具有的潜力。中国科学院化学所竺安教授也很早从事此项研究。自 1989 年有商品化仪器供应后,毛细管电泳的应用研究更是日新月异[6,7],开始主要用于蛋白质和多肽方面,以后应用在核酸[8,9]、糖[10,11]、维生素、药品检验、无机离子、环保等各种领域,和 HPLC 一起成为分析方法中互补的技术。

我们从 1990 年开始引用这项技术[12~14],继后发展用在酶活性测定[15]、磷酸化分析[16]、糖研究[17~21]中,并成功地将毛细管电泳与质谱技术联用[22~25]。本章仅就毛细管电泳概念和我们在应用中的体会作一介绍。

2. 基本原理

(1) 毛细管电泳

所谓毛细管电泳是在内径为 25～100μm 的弹性石英毛细管中进行电泳,毛细管中填充了缓冲液或凝胶。与平板凝胶电泳相比,毛细管电泳可以减少焦耳热的产生,这主要是

由于毛细管内径细,因此表面积与体积比大,易于扩散热量。另外,电泳时毛细管的高电阻限制了电流的产生和管内发热,选用高电压(10～30kV)产生高电场强度(100～500V/cm),但仍可维持较小的电流。毛细管电泳通常在高电场下进行,可以缩短分析时间并且提高分辨率,其理论塔板系数可高达 10^6。电泳所需样品体积很小(1～50nl 的进样体积),且多种分离模式能改变选择性,扩大应用范围。

(2) 毛细管电泳仪

目前生产毛细管电泳仪的厂商不下数十家,其基本结构都是由毛细管、高压电源、电极及电极液、在线检测器、恒温装置、样品盘、数据收集和处理系统等组成,如图 7.1 所示。

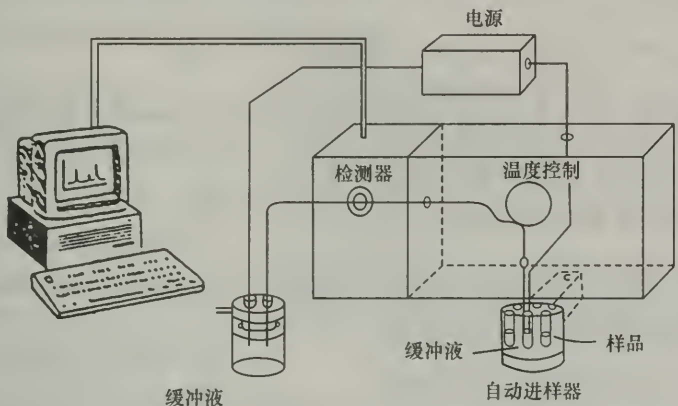


图 7.1　毛细管电泳仪结构图

毛细管是由熔融石英(fused silica)制成的,内径通常为 25～75μm,外径为 350～400μm,在其外壁涂有一层聚亚胺(polyimide)保护层,使得毛细管有一定柔性不易折断。在检测窗口需将这层保护材料去除(可用小火灼烧几毫米,再用乙醇将黑灰抹去)。去除了保护层的毛细管非常脆弱,容易折断,操作时要格外小心。

毛细管电泳仪的高压装置一般可允许电压高至 30kV,最大电流为 200～300mA,并且可改变其正、负极方向。

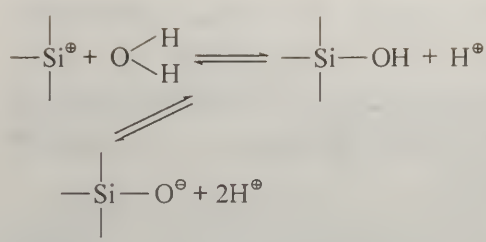
在电泳过程中,毛细管两端及电极插至电极液中,此电极液通常与管中的缓冲液一致。正、负两电极连接至高电压装置。进样时样品盘移动,使毛细管的进样端及此端的电极(通常是正极)准确插入样品管中,给予一定电场或压力后,样品被吸入毛细管内,然后再将毛细管移至电极液中开始电泳。

常用的检测方式是紫外-可见光吸收。检测器位于距样品盘约毛细管总长的 2/3 至 4/5 处,对毛细管壁内部分进行光聚焦。此外,荧光、激光诱导荧光、质谱等检测方法也被应用于毛细管电泳。

虽然毛细管的内径很细,有助于散热,但是由于 1℃ 的温度变化会引起缓冲液黏度 2%～3%的差异,因此仍需借助有效的散热装置,如简单的风扇以及含有致冷液体的冷却系统。理论上,液体冷却的效果更好,但如果空气流动速度达到 10m/s 时,也足以能够将电泳产生的热量散去,使毛细管周围的温度调节至±0.1℃的范围内。

(3) 电渗流

毛细管电泳操作中一个必需考虑的因素是电渗流(EOF)。电渗流是毛细管内壁表面



电荷所引起的管内液体的整体流动，来源于外加电场对管壁溶液双电层的作用，它通过对溶质淌度叠加一个体相流速来控制溶质在毛细管内停留的时间。

在水溶液中大多数固体表面都带有多余(额外)的负电荷，这是由于酸-碱平衡造成的离子化和(或)固体表面吸附了离子基团。构成毛细管的熔融石英由硅醇(—SiOH)组成，它在水溶液中以—SiO^-—形式存在(—SiOH的解离常数约为 pH 1.5)。

为了维持电荷平衡，溶液中的正离子吸附至石英表面形成双电子层，当在毛细管两端施加了电压后，这一层正离子趋向负极移动，并带动毛细管中的溶液以液流(bulk solution)的形式移向负极(图7.2)，这一现象称为电渗(electroosmosis)或电渗流(electroendosmotic flow，EOF)。由于毛细管的表面积与体积之比大，且电泳时使用了高电压，电渗流在毛细管电泳中常常起着不可忽视的作用。

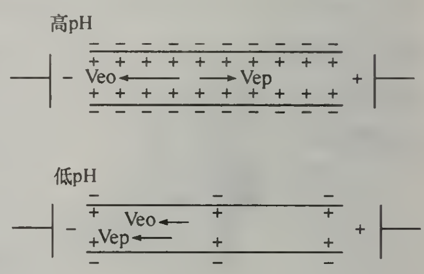


图 7.2　电渗流示意图
其中 Veo 为电渗流方向，Vep 为样品分子泳动方向

电渗流的大小可以下面两个公式表示：

$$V_{EOF} = (\varepsilon\zeta/\eta)E$$

$$\mu_{EOF} = \varepsilon\zeta/\eta$$

式中 V_{EOF}为电渗流速率；μ_{EOF}为电渗流泳动力；ζ 为电势能；ε 为双电子常数；η 为溶液的黏度。

ζ 势能是由毛细管壁的表面电荷决定的，而表面电荷依赖于溶液的 pH，因此电渗流和电泳缓冲液的 pH 密切相关。在高 pH 环境中硅醇基团易于去质子化，电渗流大于低 pH 溶液。

ζ 势能和缓冲液的离子强度也有关，增大溶液的离子强度即减少 ζ 势能，也就降低了电渗作用。

电渗作用特点之一是液体沿着毛细管壁均匀流动，因而其前沿是平头的，不同于高效液相色谱中泵推动的液体流动呈圆弧状。

另一特点是电渗使得携带不同电荷的分子朝一个方向移动，不带电荷的中性分子也能随着电渗流一起移动。一般情况下，电渗流的方向是由正极到负极。

电渗的上述两个优点在毛细管电泳分离中起着有益的作用，但是在实验中有时能够发现选择过高 pH 缓冲液进行电泳时电渗流太快以至不能将被分析物分开；而选择过低 pH 缓冲液时，毛细管内壁的负电荷也能吸附被分析物中带正电荷的分子。

测定电渗流的大小可采用中性标准物质，如二甲基亚砜(DMSO)、氧化甲基硅(mesityloxide)以及丙酮。

3. 影响因素

在毛细管电泳中，分离效率的好坏取决于区带的分散情况，而引起区带分散的因素很

多，最重要的有焦耳热引起的温度梯度、进样长度、溶质与毛细管壁的相互作用、样品和缓冲液电导不匹配等。

(1) 焦耳热

在毛细管内进行电泳的一个主要的好处是减小了曾限制传统电泳技术的热效应。热效应可导致不均匀的温度梯度和局部的黏度变化，从而引起区带变宽。尽管在毛细管电泳中高电场强度提供高分离效率，但它所带来的收获最终将受到焦耳热的限制。

电流作用下，温度的上升与产生的功率有关，并由毛细管尺寸、溶液的电导和外加电压所决定。因此，降低电场强度，减小毛细管内径，降低缓冲液浓度和离子强度，主动控温等方法都可有效地减小温度梯度，但同时可能损失一定的分离效率。

(2) 进样长度

在进样过程中，减小样品长度(进样量)非常重要，如其长度超过因扩散造成的区带分散，分离效率将受影响。在实际操作中，对进样长度的限制是低于毛细管总长度的 1%～2%。对于 70cm 长的毛细管，1%的长度相当于 7mm(或 14nl，50μm 内径)。

(3) 溶质与毛细管壁的相互作用

溶质与毛细管壁之间的相互作用可能会出现拖尾峰甚至发生对溶质的完全吸附，引起吸附的主要原因是阳离子溶质与负电表面之间的离子相互作用和疏水相互作用，对多肽和蛋白质来说这种吸附作用特别严重。

可以采取多种方法来减小溶质与管壁的相互作用。如增加缓冲液浓度可以降低有效表面电荷来减小溶质与管壁的相互作用；选择高离子强度溶液，尤其是两性离子缓冲体系。限制吸附的另一尝试是在极端 pH 下进行分离。在低 pH 下(pH 2～3)，电渗流接近于零，但因蛋白质可以质子化而带正电荷，它们仍可向负极迁移；正高 pH 下(pH 为 9～10)，毛细管壁和溶质都将解离而带负电，溶质与管壁的相互作用则因电荷之间的排斥受到限制。对毛细管进行涂层处理，通过降低相互作用的自由能来抑制吸附，也是一种有效的方法。

(4) 电分散作用

样品和缓冲液电导不匹配时会引起各区带中电场强度的变化，从而导致了峰形畸变，这种现象为电分散作用。当样品电导高于缓冲液电导时，样品的前沿将会扩散而后沿尖锐；相反，当样品电导高于缓冲液电导时，前沿尖锐而后沿扩散；当两者电导相同时，峰形对称。峰形畸变这种现象只能通过缓冲液和样品电导之间的匹配才能消除，但在分离度不受影响情况下是可以忽略的。

二、材　　料

(1) 设备和材料

①毛细管电泳仪；②毛细管；③脱气装置(水泵)；④移液器；⑤Eppendorf 管，Tip；⑥离心机；⑦pH 计；⑧过滤器；⑨微孔膜(0.22μm)；⑩针筒(10ml)。

(2) 试剂

①氢氧化钠；②盐酸；③磷酸氢二钠；④磷酸二氢钠；⑤氯化钠；⑥乙酸钠。

(3) 蛋白质和多肽

①多肽 Kemptide;②蛋白激酶 A;③巨噬细胞集落刺激因子(GM-CSF);④核糖核酸酶 T1;⑤α-乳清蛋白;⑥胰凝乳蛋白酶原 A;⑦溶菌酶;⑧α-乳白蛋白;⑨碳酸酐酶;⑩卵白蛋白;⑪牛血清白蛋白;⑫碳酸酐酶 b;⑬β-半乳糖苷酶;⑭肌球蛋白;⑮肠促胰酶肽片段。

三、基本操作步骤

毛细管电泳的基本步骤包括清洗毛细管、更换电泳缓冲液、进样、电泳并同时在线检测。

(1) 清洗毛细管

对于一根新的或久未使用的毛细管,需用 1mol/L 的 NaOH 溶液、0.1mol/L NaOH 溶液、超纯水依次清洗。在某些情况下还需用 0.1mol/L HCl、甲醇或去垢剂清洗。强碱溶液可以清除吸附在毛细管内壁的油脂、蛋白质等,强酸溶液可以清除一些金属或金属离子,甲醇、去垢剂可去除疏水性强的杂质。

每次在进行分析前可用相当于毛细管总体积 2～3 倍的 0.1mol/L 的 NaOH 溶液清洗一遍,一般为 2～3min,而后注入电泳缓冲液,以保证分析结果的重复性。

(2) 更换电泳缓冲液

清洗液、电极液(电泳缓冲液)、样品均置于可旋转的进样盘中。按照设定的指令,毛细管电泳仪可自动转动进样盘,将一端(通常是正极端)的毛细管和电极插入电泳缓冲液中,并将缓冲液吸入或推进至毛细管中。上一步清洗过程结束后,毛细管和电极从清洗液中移至电泳缓冲液,不可避免地将强碱或强酸等溶液带至其中。吸取样品后,毛细管外壁黏附的样品液也会污染电泳缓冲液。缓冲液本身置于样品盘中也会因挥发而改变其离子强度。因此,对于精确分析,每分析 5 次后需更换一次样品盘中的缓冲液。一般分析时,半天更换一次即可。

有些品牌的毛细管电泳仪具有自动更换电泳缓冲液的功能,即将进样盘中已用过的缓冲液排入废液瓶中,并从贮存瓶中引入新鲜的溶液。

(3) 进样

有 4 种进样方式。

第 1 种在毛细管进样端施加一定压力推入样品。

第 2 种将出口端连接真空装置吸取样品。这两种进样方式称为流体静力(hydrodynamic)方式,其优点是进样体积和样品浓度、样品溶液离子强度等基本无影响。

第 3 种方式为电动(electrokinetic)进样,即在一定的电场下,样品通过电渗流及自身的泳动力进入毛细管中。这种方式的上样量受电渗、样品浓度、样品分子的泳动力以及样品溶液中无机离子的多少及种类等多种因素影响,因此其进样的重复性不如前者。但是,当毛细管中填充了黏性较大的介质或凝胶时,只能选择电动进样方式。

第 4 种为虹吸,即依靠毛细管进样口端与出口端位置的高低吸入样品。

无论采用何种进样方式,毛细管插入样品溶液的深度一般要低于毛细管总长度的 1%或 2%。以尽量减少样品溶液经毛细吸附进入毛细管,从而影响进样量的精确性;样

品管中溶液最少需 5μl，进样体积为 1～50nl。因此，每次分析所耗的样品量只及其 1/1 000。样品可多次分析，或经毛细管电泳后再进行 HPLC 分离、序列测定等工作。

为防止样品挥发造成浓度改变，可对样品管加橡皮盖，盖子中间留有一缝隙，毛细管和电极可以自由出入。将样品盘连接至外围的冷却装置也可减少样品的挥发，并且还可保持蛋白质的活性。

虽然绝大多数仪器进样时间可精确至 0.1s，但是由于毛细管在插入样品管时不可避免地会黏附一些溶液，进样时间过短会影响分析结果的重复性。通常进样时间可选择 0.5～1s，此时黏附的样品量相对于应吸入的体积可以忽略不计。

受毛细管内总体积的制约，对于浓度很稀的样品，毛细管电泳不能像 HPLC 那样可以加大进样量来提高检测灵敏度，但是可以采取一些措施得以改善。

首先可以通过样品缓冲液和电泳缓冲液离子强度的不同来实现。当样品缓冲液的电导明显低于（100 倍以上）电泳缓冲液的电导时，一旦在毛细管两端通上电压，可在样品区带前形成一个更大的电场，使得样品溶液中的分子迁移更快。当这些分子趋向于到达电泳缓冲液边界时，这个附加产生的电场减弱，这些分子迁移减慢，直至样品移至电泳缓冲液边界，样品区带内电场均一。样品区带由宽变窄，样品中各成分得以浓缩。

假如样品缓冲液和电极缓冲液的电导相同，而样品浓度较稀，不适合再作稀释，可以在进样前先吸入一些水，也能达到浓缩的目的。无论采取何种进样方式，上述方法都可以使样品自动堆积浓缩。

必须注意的是，采取这种浓缩堆积方法时，由于在堆积区带电势陡降，会导致此区带温度升高。

另外，有报道将若干根毛细管平行排列成毛细管束进行电泳以提高分析灵敏度[26]，不过此法需在改进的仪器上进行。

(4) 电泳和检测

石英毛细管可以透过 190～700nm 范围内的光，在实验时应尽量选用低波长检测以提高灵敏度。与此相应，电泳缓冲液必需在低紫外波长下的吸收值低，否则会增加基线噪声并降低检测信号。磷酸盐、硼酸盐等无机盐缓冲液符合上述条件，因而被广泛使用。选用生物实验中常见的 HEPES、CAPS、Tris 等缓冲液时检测波长不得低于 215nm。

采用荧光或激光诱导荧光检测，非常灵敏，质量检测限最高可达 10^{-20}mol，但常常需要样品衍生化。

在分析稀浓度样品时，也可以通过扩展检测光程来提高灵敏度。具体方法是安装在检测窗口处呈球泡状的毛细管，或在此处加工成“Z”字形结构。

与检测器相连的记录系统（记录仪、积分仪、电脑等），除可显示分离图谱外，还可以根据峰面积积分进行定量分析。但是，和 HPLC 不同，不同的分子在毛细管电泳中的泳动速率不一致导致泳动慢的分子积分面积大，这时可以将峰面积除以迁移时间进行校正。

四、毛细管电泳分离模式及应用

与 HPLC 相似，毛细管电泳根据不同的分离机制，有多种分离模式，其中在蛋白质化学研究中经常应用的模式见表 7.1。

表 7.1　毛细管电泳分离模式

模式	分离机制
毛细管区带电泳(CZE)	荷质比差异
微团电动毛细管色谱(MECC)	与微团的疏水/离子交联
毛细管筛分电泳(CSE)	大小与电荷
毛细管等电聚焦(CIEF)	等电点

下面分别就上述模式的特点及其应用进行介绍。

1. 毛细管区带电泳(capillary zone electrophoresis, CZE)**或称自由溶液毛细管电泳**(free solution capillary electrophoresis, FSCE)

毛细管中充满了电泳缓冲液后,由于电渗的作用,分析物中带正、负电荷及中性的分子均能向一个方向移动,但是不同的中性分子没有自身的泳动力,都和电渗流共迁移,它们之间不能分开。

由于 CZE 具有与反相 HPLC 不同的分离机理,且其理论塔板系数不亚于后者,可成为后者的一项互补技术,尤其对于亲水多肽的分离分析具有一定的优越性。在某种程度 CZE 较 HPLC 更方便,有更多的选择性,只需改变电泳缓冲液的条件,(如 pH、离子强度等),不必更换毛细管。在合适的电泳条件下,毛细管电泳可以分辨多肽的细微结构差异,如 C 端的酰胺化与去酰胺化,单个氨基酸残基的替换,氨基酸组成相同而序列不同的多肽也能区别开,甚至表面电荷不同而一级结构相同的多肽也有可能分离开。还有报道用 FSCE 帮助确定多肽内是否存在二硫键[3]。

图 7.3 为多肽 kemptide 经蛋白激酶 A 作用后的磷酸化产物与底物的分离图谱。由于毛细管电泳分析快速,我们以此建立了测定蛋白激酶 A 活性的方法,其灵敏度、精确性和线性范围均不亚于经典的同位素法[15]。

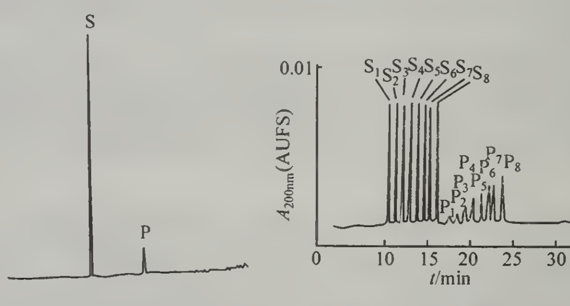


图 7.3　多肽 kemptide 与其磷酸化产物的分离图谱
S 为底物 kemptide, P 为磷酸化产物

在运用 CZE 这一模式进行分析时首先要考虑的是如何选择缓冲液,一般掌握的原则是:①在所选择的 pH 范围内有很强的缓冲能力;②在所选择的检测波长缓冲液本底吸收低;③溶液中的离子泳动力低(即结构大、带电荷少),从而可以减少电流。

表 7.2 为在 CZE 中常用的缓冲液及其 pK_a 值,可供参考。

表 7.2　常用缓冲液及其 pK_a 值

磷酸盐	2.12(pK_{a1})	MOPS	7.20
柠檬酸盐	3.06(pK_{a1})	磷酸盐	7.21(pK_{a2})
甲酸盐	3.75	TES	7.50
琥珀酸盐(succinate)	4.19(pK_{a1})	HEPES	7.55
柠檬酸盐	4.74(pK_{a2})	HEPPS	8.00
乙酸盐	4.75	Tricine	8.15
柠檬酸盐	5.40(pK_{a3})	甘氨酰甘氨酸	8.25
琥珀酸盐	5.57(pK_{a2})	Tris	8.30
MES	6.15	Bicine	8.35
ADA	6.60	Morpholine	8.49
BIS-TRIS 丙烷	6.80	硼酸盐	9.24
PIPES	6.80	CHES	9.50
ACES	6.90	CHAPSO	9.60
MOPSO	6.90	CAPS	10.40
Imidazole	7.00	磷酸盐	12.32(pK_{a3})

图 7.4 为用不同 pH 的缓冲液鉴定遗传工程重组蛋白 GM-CSF 的纯度。可以发现在 pH7.0 的磷酸缓冲液中样品的不均一性更加明显。由此建议,用毛细管电泳鉴定遗传工程产品时,宜选择多种缓冲液进行分析。一般选在蛋白质 pI 的两侧。

选择合适的温度也是应用 CZE 时应该注意的,这不仅是因为温度的高低可以改变缓冲液的黏度、电渗以及分析时间,而且温度将影响蛋白质的构象、蛋白质与辅基的交联等。

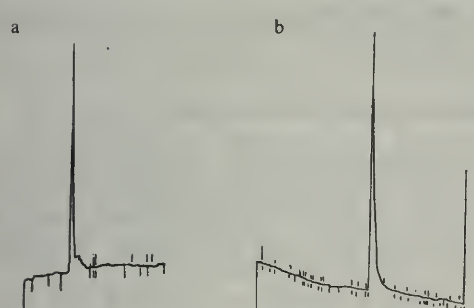


图 7.4　用 CE 鉴定 GM-CSF 的纯度
a 为 20mmol/L 磷酸缓冲液,pH7.0;
b 为 50mmol/L 磷酸缓冲液 + 5mmol/L NaCl,pH2.5。
电压均为 30kV,进样 0.5s,检测波长 214nm,毛细管长度 72cm(至检测处有效长度 50cm),内径 50μm

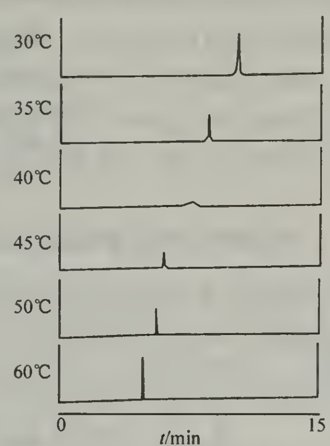


图 7.5　核糖核酸酶 T_1在不同温度下的毛细管电泳结果
电泳条件:25mmol/L 磷酸钠缓冲液,pH7.0;电压 15kV,毛细管长 72cm,有效长度 50cm,内径 50μm,检测波长 200nm

图 7.5 显示不同温度对核糖核酸酶 T_1在毛细管中电泳行为的影响。随着温度的升高,核糖核酸酶 n 在毛细管中泳动速度加大,但在 40℃此蛋白发生了构象变化(或称热变性)。

对于等电点偏碱性的蛋白质(例如细胞色素 c),在一般的电泳缓冲液中带有较多的正电荷,容易被毛细管内壁吸附。解决这一问题的方法之一是对毛细管内壁进行涂层以

改变电荷极性或使毛细管内壁电荷趋于零。目前有商品化的内壁经过共价永久修饰的毛细管,这类毛细管的内壁几乎不带电荷,如J&W公司提供的DB-1毛细管。此外,Supelco公司有不同程度亲水性质(P1系列)和疏水性质的(H、H1、H2系列,分别为C1、C8、C18)的毛细管供应。它们的优点是:可以减少样品与毛细管壁的相互作用;减小电渗流;增加重复性和选择性。另外也可在实验室对毛细管进行动态可逆修饰,比如Perkin-Elmer公司的产品Micro-Coat试剂盒即可对普通的石英毛细管进行涂层使内壁带正电荷,如不需要涂层,用1mol/L的NaOH溶液即可将此涂层除去。

图7.6为我实验室进行的三种蛋白质涂层前后毛细管电泳图谱的比较。

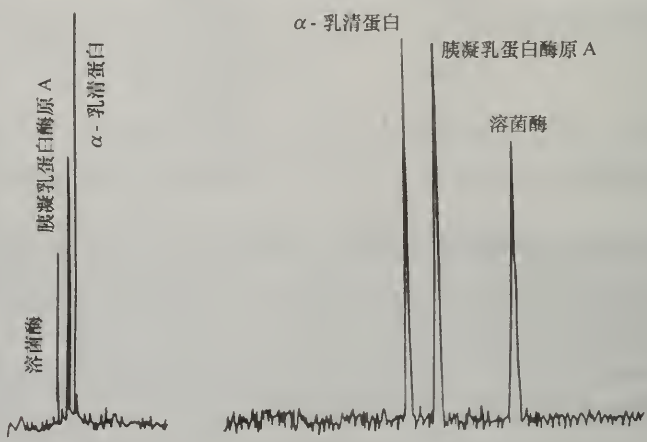


图7.6　毛细管涂层前后分离蛋白质的比较

电泳条件:40mmol/L乙酸钠,pH4.0,真空进样2s,

200nm检测涂层前25kV,涂层后-25kV

2. 微团电动毛细管色谱(micellar electrokinetic capillary chromatography,MECC)

MECC在某种程度很像反相HPLC,通常在电泳缓冲液中添加一定浓度的表面活性剂,这类分子一端为亲水基团,其余部分为疏水结构。当表面活性剂的浓度高于其临界微团浓度(critical micelle concentration,CMC)时,表面活性剂分子就会发生聚集,形成微团或胶束(micelle)。其中疏水部分在内部,亲水基团在表面,被称为假固相。在毛细管内,样品中的各种分子除了电泳外,还与这些微团存在程度不同的吸附。根据在假固相和溶液相中的分配不同各种分子得到分离。前面提及的在FSCE中不能分开的中性分子可以经MECC分析。

表7.3列出了常用的表面活性剂,其中最常用的是SDS。

表7.3　常用的表面活性剂

类型	名称	浓度/(mmol/L)	聚合数目
阴离子	SDS	8.2	62
阳离子	DTAB	14	50
	CTAB	1.3	78
非离子	Triton X-100	0.24	140
两性离子	CHAPS	8	10
	CHAPSO	8	11
胆汁盐	cholic acid	14	2～4
	deoxycholic acid	5	4～10
	taurocholic acid	10～15	4

采用 SDS 构成假固相时，微团带有大量的负电荷，其泳动方向与电渗相反，但是如果选择 pH 偏碱的缓冲液可维持电渗流速率大于微团的泳动速率，因此微团的最终泳动方向仍与电渗流一致。样品中的中性分子疏水性越强，与微团的结合越紧密，迁移时间也就越长。其他带电荷分子除了与微团形成疏水吸附，还有离子交联等作用，最终泳动速率取决于其疏水程度、带电荷多少等各种特性。

在应用 MECC 模式时常可加入有机修饰剂，如甲醇、乙腈、异丙醇等。这些有机溶剂可以降低溶质和微团间的疏水吸附，也可减小形成微团的疏水作用。

由于 MECC 可以同时分离带电荷与不带电荷的分子，目前被较多地应用于氨基酸衍生物、肽谱分析以及合成多肽纯度鉴定等。

3. 毛细管筛分电泳(capillary sieving electrophoresis，CSE)

此种电泳模式的特点是毛细管内填充了多聚物作为分离介质，对分子的泳动起着阻碍作用。分离介质可以如平板电泳一样呈凝胶状，也可以是甲基纤维素溶液，后者便于清洗，可实现毛细管一管多用。溶液中的甲基纤维素分子在电子显微镜中呈现缠绕状，其作用机理类似于凝胶的筛网作用，大分子较小分子所受的阻碍大，泳动速度减慢。这种分离模式尤其适合于荷质比差异很小而大小不同分子的分离(如 DNA 分子、SDS-饱和的蛋白质分子)。这类分子应用 FSCE 无法进行分析。

CSE 在 DNA 研究中有着广泛的应用，在蛋白质研究中主要用来代替传统的平板 SDS-聚丙烯酰胺凝胶电泳。其优点在于：第一，缩短分析时间，通常凝胶电泳至少需要 45min 左右，而 CSE 只需 15min 即可达到同样的分离效果；第二，在毛细管电泳中可以在线检测并进行更精确的定量；第三，毛细管电泳仪自动化程度高，可以省却手工灌胶、聚胶、上样、染色、脱色等操作步骤。但是 CSE 每次电泳只能分析一个样品，不能像平板凝胶电泳那样具有多个进样孔可以同步电泳。另外，目前 CSE 只能用于相对分子质量 1 万至 20 万范围内蛋白质的分离，尚不可进行多肽的分析。

图 7.7 为 7 种相对分子质量不同的标准蛋白质经 CSE 的分离图谱。

应用毛细管 SDS 凝胶电泳可以比较方便地分析糖蛋白的分子质量。由于多糖基团不结合 SDS，糖蛋白的荷质比相对于同样大小的蛋白质低，用经典的 SDS-聚丙烯酰胺凝胶电泳估计的分子质量偏大，通常必需在不同浓度的凝胶上进行分析，得到糖蛋白迁移率对数值相对于凝胶浓度的线性关系图(Ferguson 图)，该直线的斜率称为此蛋白的排阻系数 K_r。K_r 和蛋白质的半径的平方成正比，而蛋白质半径与其分子质量的对数成线性。因此，得到标准蛋白分子质量的对数值相对于 K_r 平方根的线性关系图，测得糖蛋白的 K_r 值即可推测其分子质量。运用传统的平板凝胶电泳需多次灌胶多次电泳，而毛细管凝胶电泳应用溶液状的分离介质只需对浓的含甲基纤维素的分离溶液进行一系列的稀释，按照设定的程序自动分析即可得到 Ferguson 图。

图 7.8a 为用常规方法测定糖蛋白的分子质量，可发现根据相对迁移率推测的糖蛋白分子质量将与实际数值有较大差异。

图 7.8b 为用 Ferguson 分析方法测定糖蛋白的分子质量，图中显示糖蛋白与标准蛋白基本上位于一条直线上，也就是说根据 K_r 的平方根值，可以较准确地确定糖蛋白的分子质量。

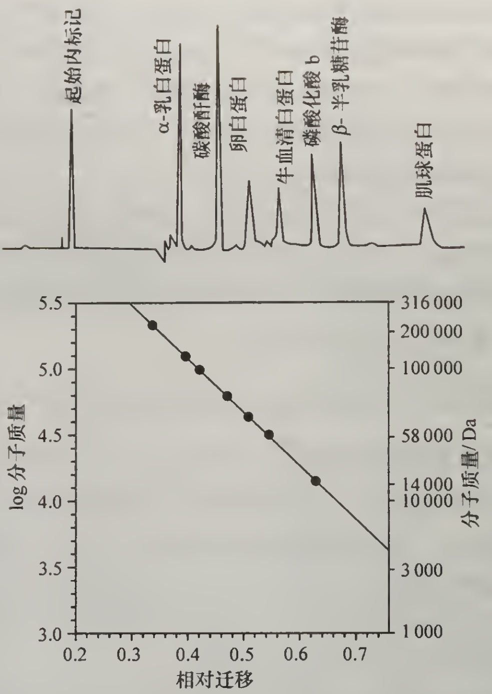


图 7.7　标准蛋白质的 CSE 分析

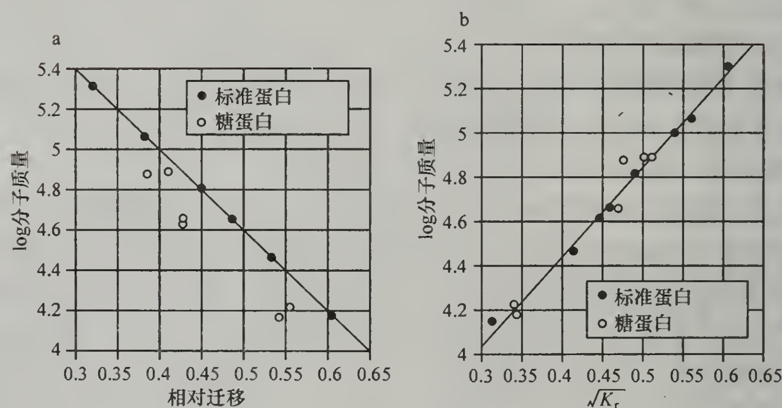


图 7.8　糖蛋白的分子质量分析

a. 一种凝胶浓度测得的相对迁移率；

b. 用几种凝胶浓度进行 Ferguson 分析

4. 毛细管等电聚焦(capillary isoelectric focusing, CIEF)

与普通的凝胶等电聚焦一样，毛细管内的 pH 梯度也是由两性电解质形成。在实验中需选用内壁中性共价涂层的毛细管，阳极端至检测器有效分离部分(离检测窗口差几厘米)充满两性电解质溶液，样品溶液夹在其中以避免直接接触阳极液(H_3PO_4)。毛细管其余部分(检测器至阴极)为阴极液(NaOH 溶液)。毛细管两端分别插入阳极和阴极液中，其中毛

细管中的溶液和 H_3PO_4 溶液中均含一定浓度的可溶性甲基纤维素作为支持介质。两性电解质载体和样品在电场中聚焦至电流趋于零。此时可采用真空抽吸、压力推动或者在电极液中加入盐以改变已经形成的 pH 梯度等方式驱使蛋白质条带移动并逐一经过检测窗口。

需要注意的是,毛细管等电聚焦必须选用内壁经过共价涂层消除了电渗的毛细管。与前面介绍的 CSE 一样,CIEF 不仅可以取代常规的凝胶等电聚焦,还具有前述的诸多优点。另外,由于是直接在线检测,CIEF 可以分析多肽的等电点,而在传统凝胶等电聚焦中的固定、染色、脱色等步骤中小分子多肽易丢失。

整个分析过程可自动化,通常仅需 45min 左右。图 7.9 说明毛细管等电聚焦的过程。

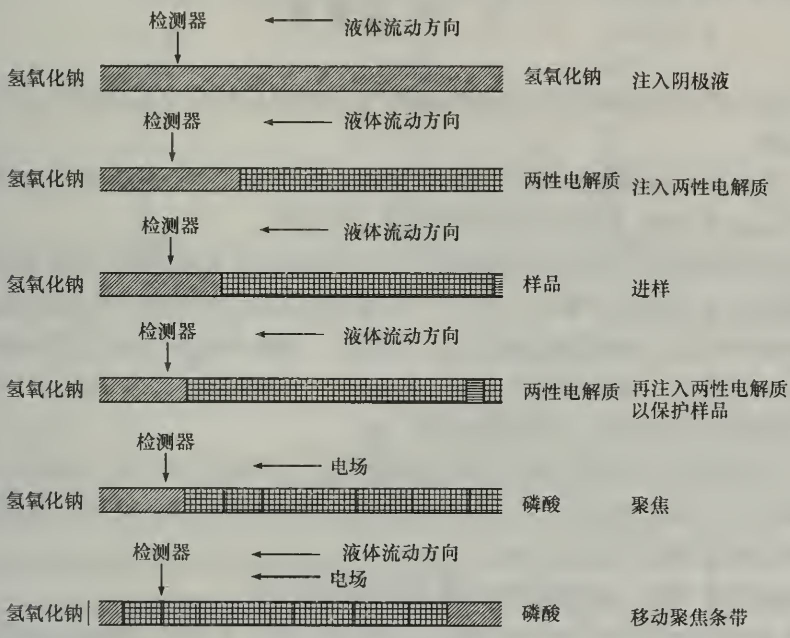


图 7.9　CIEF 程序

图 7.10 为我们实验室应用 CIEF 测定重组人表皮生长因子(rhEGF)的等电点,经分析其 pI 为 5.0[27]。

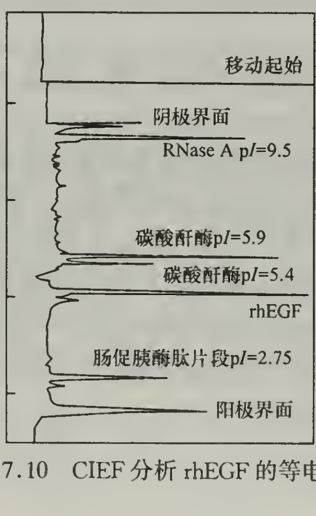


图 7.10　CIEF 分析 rhEGF 的等电点

由于等电点与聚焦位置距离检测窗口的长度呈良好的线性关系，用 CIEF 可以测到以往用凝胶 IEF 较难测准的接近 pH 梯度极端值的等电点，如 pI9.0 以上或 pI3.0 以下的蛋白质[28]。

另外，CIEF 具有较高的分辨率，可分辨至 pI 相差 0.5 个单位，因而它可用于分离差异较小的异构体。

CIEF 还具备其他毛细管电泳分离模式没有的优点，即对于较稀蛋白溶液（低至 0.001μg/μl）可将它们和两性电解质混合后一起注入毛细管中进行分析。

五、注意事项

①毛细管的清洗和平衡条件应保持一致，尤其在做多样品定量分析时是保证高重现性的重要因素。

②用于毛细管电泳的缓冲液应尽量新鲜配制，使用的试剂为电泳纯，水的质量要求也很高，以 Milli-Q 水为佳，使用前还应脱气和膜过滤处理。

③正负极两端的缓冲液必须是相同的，如不同，会使吸光值增减，电流变化以及电渗流改变。两端的缓冲液池要等高，这样对于分离效率和迁移时间重现性都是很重要的。

④缓冲液的更新是获得高重现性的一个环节。溶液的电解会改变缓冲液的 pH 值，电渗流也会随之改变。电解的程度决定于电流的大小和分析时间的长短，pH 值变化的程度决定于缓冲液的缓冲容量，两端池体积等。

⑤用于毛细管电泳的样品应均匀地呈溶液状态，上样前应离心取上清分析。

⑥如果没有灵敏度的限制，应该用尽可能短的进样长度，但是受到仪器的限制，进样时间短会使进样重现性减小。

⑦为防止样品蒸发，必须盖紧样品管盖或冷却样品盘。

⑧为防止样品沾带引起额外的进样，毛细管端口要平正光滑，并除去毛细管端口的聚酰亚胺。另外，进样后要及时将毛细管浸入缓冲液或水中。

⑨毛细管电泳结束后，如有一段时间不用，应将毛细管作清洗后，以氮气吹干管内留有的液体后保存。

六、结　束　语

毛细管电泳是基于电泳的高分辨机制和色谱的仪器自动化概念而诞生的，因而兼具二者的优越性，即分辨率高、分析速度快、消耗样品少、灵敏度高、操作简便、分离模式多样、不使用毒性强的有机溶剂等诸多优点。表 7.4 显示毛细管电泳与其他分析方法的比较结果。毛细管电泳技术发展迅速，具有很好的应用前景。但是，它只能局限于分析水平，尚不能作为分离制备的手段，即使用多次分离得到的样品，仅能进行如 N 端序列、质谱等微量分析，而且它还不能分析难溶于水的样品，因此不能完全替代气相色谱、液相色谱及传统的凝胶电泳。

表 7.4　毛细管电泳与其他分析方法的比较

项　目	毛细管电泳	凝胶电泳	HPLC
分离机制	电荷、疏水性分子大小等	大小与电荷	极性/非极性；分配大小；离子交换
常规分析时间	10～30min	2h～几天	10～120min
分辨相近分子的能力	优秀	优秀	良好
样品需要量	<1μl	1～100μl	10～50μl
灵敏度(一般常规分析)	皮克至纳克	纳克至微克	纳克至微克
定量准确性	好	较差	好
检测方式	在线检测可见-紫外吸收荧光，激光	染色，放射自显影	在线检测可见-紫外吸收荧光，激光，电化学等
仪器价格	较高	较低	较高
实验消耗	低	较高	较低
手工操作	少	多	少
自动化程度	高	低	高
制备能力	无	有	高
部分收集	困难	困难	容易

第二节　连续自由流电泳

一、概　　况

1. 概述

Philpot 在 20 世纪 30 年代后期首先考虑到连续电泳的优点，至 10 年后 Svenson 和 Brattsten，Grassmann 以及 Hannig 便报道了在这方面所做的首次实验，他们研究了被电解液流所带动的颗粒在经过垂直于流动方向的电场作用后其区带的迁移原理。然而，因系统本身的复杂性，以及对一些基本现象缺乏认识和仪器条件的限制而未被广泛应用。直至 80 年代后，科学家们又考虑到液相电泳的特点，相继发展了毛细管电泳和制备级的连续自由流电泳、液相等电聚焦，后两者使电泳的大规模制备型分离得以可能。

连续自由流电泳(continuous free flow electrophoresis，CFFE)可理解为在无固相支持介质的薄的矩形分离腔中，以缓冲液作为分离介质进行生物大分子或不溶性颗粒及细胞等物质的电泳分离纯化技术，它可采用多种类型的仪器系统。也被称为 Hannig 型自由流电泳，简称自由流电泳。曾有不少学者对自由流电泳进行过全面的综述[29~33]。

自由流电泳是基于区带电泳的原理——当带电样品以细带进入系统后，在电场的作用下根据它们的电泳迁移率在一定 pH 介质中得以分离。从理论上看，这种过程较等电聚焦更灵活，因为等电点相似的样品在另一 pH 介质中能有相当不同的迁移率。然而，区带电泳不可避免的次现象如分子扩散等使样品带变宽因而导致分辨率下降。所以，最有利的是提供短而快速的分离，而自由流电泳正是满足了在电场中滞留时间短的要求。另外，更引人注目的是，它除了能分离离子和大分子物质如蛋白质等外，还能分离大颗粒甚至活细胞，这是其他过程所不具备的。

对比其他分离技术而言，自由流电泳的主要优点在于它是一个连续的而不是分批进

行的分离过程，同时，由于它一般不使用有机溶剂、高盐溶液及硅胶或凝胶等支持介质，其分离条件相当温和，对有活性的生物材料的分离纯化提供了十分合适的分离环境。然而因自由流电泳是完全无载体的液相电泳，因此除了电泳本身所固有的影响因素（焦耳热，电动力学变形）外，还有层流（流体力学变形）以及一些综合因素的影响（电流体力学变形等），且这些因素常常相互关联，使整个过程变得极为复杂。也许正因为此，自由流电泳在很长时间内被忽视。然而，尽管众多因素会引起样品流的扰动而降低自由流电泳的分辨率，使其应用受到限制，但在大规模分离的可能性以及多种生物样品（如离子、大分子、颗粒、完整细胞、细胞碎片等）分离的可行性上显示了独特的优点，再加之近 20 年来空间自由流电泳的开展，使这种被忽视的局面得以改善，并推动了该领域研究的发展。

我们实验室首先选择两种有色的等电点差异较大的标准蛋白质作为分离样品[34~38]，旨在认识和研究自由流电泳的各种现象，之后进行了一些地面应用研究，主要为药用蛋白质和细胞的分离。天花粉蛋白是一种核糖体失活蛋白，对其粗产品的自由流电泳分离，得到经多种分析技术鉴定为纯度很高（大于 99%）的产品[39]；还进行了一些基因工程粗产品如干扰素、人表皮生长因子的分离尝试。对细胞的分离是在人和兔红细胞成功分离的基础上[34]，又分离了小鼠脾脏淋巴 B 细胞和 T 细胞[40]、传代培养的肝癌 7414 细胞和鼻咽癌细胞 CNE-2，基本建立了细胞的自由流电泳分离体系。

2. 基本原理

自由流电泳的原理图见图 7.11。在 CFFE 中，分离介质在两块平行的矩形板组成的分离腔内形成层流，两板之间的距离通常介于 0.5～3.0mm（地基分离腔通常为 0.5～1.0mm 之间），分离腔的两侧为正负电极室。被分离样品由一口径很小（约为分离腔厚度 1/3）的输入口进入分离介质中，形成一细带，在垂直于液流方向上加上均匀的电场后，样品中的各组分由于各自电泳迁移率的差异而各自向与所带电荷符号相反的电极以不同的角度迁移，相同电泳迁移率的物质则迁移为一窄带，在到达分离腔出口处由级分收集器收集。

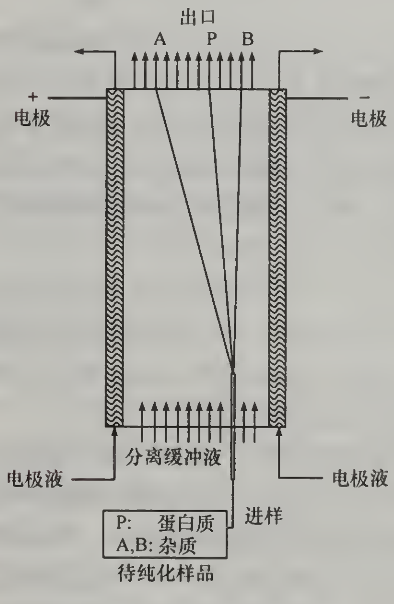


图 7.11　自由流电泳原理图

3. 主要影响因素

自由流电泳在 20 世纪 50 年代被提出后，并没有受到重视，主要的原因是其过程相当复杂，且为动态过程，导致稳定性、分辨率都不易控制，使其在很长一段时间内都没有一个可操作的模式，所以不像 SDS-PAGE、IEF 等电泳技术几乎普及于所有的生化实验室。CFFE 过程中所涉及的因素不仅种类繁多而且常常相互关联，如热对流、流体力学变形、电动力学变形、电流体动力学变形等。若把所有的影响因素都考虑在内，那么列出的方程组用现有的数学方法将无法解出。

(1) 热力学不稳定性(thermal instabilitiy)

当电流通过导电体后,由于焦耳热的产生会使导电体的温度升高。在 CFFE 中,当使用很薄(0.5～1.0mm)的分离腔时,电泳中产生的热量可以随着分离介质的高速流动被迅速带走;而当分离腔厚度增加时,处于分离腔中心部位溶液的热量不能有效地散发,从而导致温度升高。

焦耳热的产生直接导致温度梯度的形成,后者又导致密度梯度的形成。在重力加速度存在下,必然造成热对流现象,从而严重影响电泳分离效果,所以控制热对流现象极为重要。在工程上以相关无因次数(dimensionless number)Grashof 数(Gr)来描述这一现象:

$$Gr = (g \cdot \beta \cdot \Delta T \cdot r^3)/v^2$$

式中,g 为重力加速度,β 是介质的体积膨胀系数,ΔT 是介质流经分离腔后的温差,r 是分离腔的半厚度,v 是介质的黏度常数。

Gr 超过一定值后,液流的不稳定性增加,并可导致已分离组分的再混合而使分辨率降低,因此减小 Gr 是设计自由流电泳仪和实验的中心技术问题。通常,增加冷却功率、减小分离腔厚度、增加介质黏度均可使 Gr 大为降低从而保证液流的稳定性。另外,减小重力加速度也不失为一条途径,也正因为此,人们尝试在空间进行自由流电泳分离。

(2) 流体力学变形(hydrodynamic distortion)

流体力学变形是自由流电泳的另一个固有现象。

当流体在两块平行板之间流过时其速率分布会形成一种抛物线型模式,在中心部位的流体流速最大,而靠近分离腔壁的流体速度较小,在腔壁上流速为零。这就是所谓的 Posseuil 层流型速率分布。在横跨分离腔厚度方向上,液流存在着速率梯度。速率梯度造成的停留时间梯度使得各物质受电场作用的时间不同,即物质在电场方向的迁移时间不同,最终导致样品带的变宽,而且,样品带的形状将不再是笔直的而呈新月状。

为了解决这个问题,首先可以减小样品输入口的口径,这样样品流由于保持在一个距离分离腔壁较远的位置上,其中各物质的速度梯度将较小,从而具有比较均一的迁移时间,但会使电泳的物料流通量减小。另一个办法就是增加分离腔的厚度,但这同时会带来由于散热效果差而造成的热对流问题。

(3) 电动力学变形(electrodynamic distortion)

因电泳池材料和介质中阳离子引起的双电层导致的电渗流在所有液相电泳中都存在。在 CFFE 中,电渗流会把样品流变形为新月状——在靠近分离腔壁区域的离子被拖向阴极,而处于分离腔中部的离子则基本不受影响。电渗流可在一定程度上被控制和利用。当电渗流方向与物质在电场中的迁移方向相反时,能使分离的分辨率提高[29];当电渗流的速度与物质在电场中的迁移速度大小相等、方向相反时,由流体力学变形和电动力学变形所带来的影响可完全相互抵消。这就意味着必须尽量使分离腔壁上的毛电势和待分离物质的毛电势一致。用具有高亲和力且其 ξ 电势与待分离物质相近的可溶性材料来覆盖分离腔壁是一种有效而简单的方法,如在细胞分离时,分离腔壁用 2%～3% 的牛血清白蛋白处理后可减小条带扩散[41]。

(4) 电动流体力学变形(electrohydrodynamic distortion)

样品流由于蛋白质的存在常常具有与分离介质不同的离子组成,使它们之间存在电

导差异，进而导致溶液中电场的变形并在溶液中产生剪切应力，这种变形被称为电动流体力学变形，其造成的结果是样品带的变宽。当样品流比介质电导低时，发生朝向分离腔壁的变宽现象；当样品流的电导比介质高时，则产生平行于分离腔壁的条带变宽，后者常见。这种变形在自由流电泳分离过程中占有很重要的地位，而且，即使在微重力环境下这种变形效应仍可能存在。

解决这问题的简单方法是仔细地调整样品流与介质的电导，使它们大致相同[41]，但在蛋白质样品浓度高时很难做到。

(5) 电动学效应(electrokinetic/Kohlrausch effects)

一般讨论时都假设蛋白质的电泳迁移是不受周围环境影响的。但实际上在电解质溶液中加上直流电场后，溶液中的所有离子都将发生电迁移。但溶液中不同离子的迁移必须同时满足其周围的电中性和化学平衡的条件，这就意味着蛋白质样品流附近的离子浓度常常会发生不定的变化，从而产生所谓的电动学变形效应。

(6) 弛豫效应(relaxation effect)

在电解质溶液中，带电颗粒和分子都被电子云所包围。颗粒的半径 a 与电子云的伸展量 $1/\kappa$ 的比率，即 $\kappa \cdot a$，是一个可测量。当电场加上后，电子云的形状将发生变形而产生偶极移动。这种效应在电子云具有极大或极小的伸展量时都可忽略不计。而在高的 ξ 电势下，电子云具有中等的伸展量时即可达到最大的偶极移动。这将产生一种与电场力方向相反的迟滞力，这就是弛豫现象。所以，在 $1<\kappa \cdot a<100$ 时，带电颗粒的分离依据将不但有 ξ 电势而且还有颗粒的大小。这个范围内的 $\kappa \cdot a$ 值可在极低浓度的电解质溶液中达到。

可以看到，这种弛豫现象也给大型制备级分离开辟了一条新的道路。在极低浓度的电解质溶液中，物质以比在常规的较高离子强度分离介质中高得多的速率进行电迁移，所以可以应用比较高的流速而增加物料流通量。而且，由于所通过的电流强度很低，电泳过程产生的焦耳热也将减小，从而可以减轻冷却功耗。在这种低浓度电解质溶液中进行电泳，惟一可能的限制性因素就是某些样品的稳定性问题。

(7) 浮力驱动的扰动效应(buoyancy-driven flow)

分离缓冲液与样品流之间以及分离缓冲液自身之中的密度差异可以引起自然对流现象。样品流和分离介质之间有不可忽略的速率差异，而且会沿着分离腔的垂直方向随缓冲液中离子和蛋白质的不同迁移而变化，于是，蛋白质在电场中的迁移轨迹将不会像预测的那样是笔直的。同时，缓冲液自身也存在密度不均一性，这是因为缓冲液中的离子迁移受到分离腔与电极室之间分隔膜的影响，同时溶液中的离子迁移还受到由电流通过而产生的焦耳热在溶液中的扩散所带来的影响。由膜所产生的浓度极化效应还可在缓冲液内形成温度梯度，进而导致缓冲液的密度不均一性。

(8) 电极膜的影响

用于分隔电极室和分离室的电极膜，一方面可防止电极所产生的气泡进入分离腔，另一方面还可防止高浓度电极液在高速循环时对分离腔内的流场造成破坏。但是，电极膜会产生浓度极化现象，即在分离腔一侧产生一个离子积聚区而在另一侧产生一个离子耗竭区，这时、离子耗竭区由于其电导较低将比其他区域的介质温度低而形成温度梯度，进而导致介质的密度梯度，引起自然对流。

为避免膜引起的影响因素，可考虑三种方法：①为克服由经过膜的扩散性转移引起的耗竭现象，可使用薄的膜和大的电极液-分离液浓度差。②使用具相反极性的膜，并使用可在分离区域内产生两个积聚区的电流方向。合适的设置方式为：阴离子交换膜置于阴极侧，而 CEA 膜、PES 膜或其他阳离子交换膜置于阳极侧。③为控制随膜设置而变的浓度变化情况，可使用具有高浓度的边界溶液。这样，积聚区和耗竭区对分离中心区域的影响都将很小，该区域的电场强度可保持均一而稳定。

(9) 气泡

由于水水解在电极产生气泡，在负极积累 H_2，在正极积累 O_2，影响了分辨率。采用钯电极或铂黑电极，可减小这个影响，但不能消除它，特别在高电流时。在微重力下气泡更不会自动逸出，因此采用电极液循环流动去除。在分离室和电极之间用半透膜是一种有效的方法，此外还可在循环回路上加上相分离器。

另外，非常重要的是，在地基自由流电泳实验中，重力加速度是一个不可忽视的影响因素。它与许多影响因素有关，如热对流、流体力学变形、分子沉降等。这些效应在进行自由流电泳的放大实验时显得尤为严重。因此，要进一步提高自由流电泳的分辨率和产率，必须在微重力条件下进行电泳操作。不少国家已经开展了空间自由流电泳的研究工作并取得了长足的进展，对此 Snyder 和 Rhodes[42]，以及 Morrison 都作了很全面的总结[43]。

4. 自由流电泳仪

自从 Tiselius 的 U 型管电泳装置问世以来，对电泳仪器的研究已经走过了很长一段历程。可以克服对流及样品流变形等问题的方法已经得到发展，许多不同类型的仪器也已商品化。人们朝着使蛋白产率最大化、蛋白分离最佳化、自由对流不稳定性最小化以及蛋白区带扩散最小化等目标对自由流电泳分离装置的设计进行了研究和探索。

一台典型的自由流电泳仪包括四个部分：泵驱动系统、电源、级分收集器和分离单元。泵驱动系统通常采用的是一个或多个蠕动泵。输送分离介质的泵可放在分离介质入口的前面，也可放在出口的后面。分离介质泵和电极液泵可合用同一泵，也可独立使用，样品泵通常独立于介质泵和电极液泵。

分离单元包括分离介质和样品的输入口、分离腔和电极室。分离腔通常由两块玻璃或聚碳酸酯等材料制成的平行的矩形板构成，电极室设在分离腔的两侧，其中的电极由不易在电场中氧化的金属如铂制成，电极室与分离腔由电极膜如离子交换膜、醋酸纤维膜等隔开。

电源要能提供达到 300V/cm 以上的直流电压，最大电流达到 200mA 以上。

级分收集器用来收集分离腔出口处的已分开的样品和分离介质，通常由很多输出管道和接收小管或小瓶组成，也有仪器在输出管路中或管路前设置在线检测的元件。

二、材　　料

(1) 设备和材料

①自由流电泳仪（Octopus 仪，德国 Weber GmbH 生产）；②酸度计；③电导仪；④离心机；⑤磁力搅拌器；⑥分光光度仪；⑦移液器，Tip，Eppendorf 管。

(2) 试剂

①三乙醇胺;②乙酸;③甘油;④葡萄糖;⑤甘氨酸;⑥蔗糖;⑦乙酸钾;⑧氢氧化钠;⑨磷酸;⑩两性电解质(ampholine)。

(3) 样品

①牛血红蛋白;②牛血清白蛋白;③细胞色素 c;④天花粉蛋白(丙酮沉淀粗提物);⑤人和兔的红细胞。

三、自由流电泳操作步骤

自由流电泳的基本操作包括安装和清洗分离腔,缓冲液和样品的加注,电泳,收集和检测等步骤。

(1) 安装和清洗分离腔

久未使用的分离腔的前后腔壁应以酒精或异丙醇湿润的棉花球仔细擦洗。根据分离模式选择所需的电极膜,并仔细洗去电极膜上的存储液,以重蒸水润湿后与相同长度的滤纸条一起封住分离腔前壁的电极室,安装间隔条后将分离腔的前后壁合拢,并以两侧的弹簧夹固定,注意一定要用力均匀,否则腔后壁的玻璃(镜子)易破碎。

根据分离模式连接好管路及泵的设置,然后进行管路的清洗。需用 0.1mol/L NaOH、超纯水、0.1mol/L HCl、超纯水依次清洗。如管路特别脏,可以异丙醇或去垢剂清洗。在最后超纯水清洗过程中还需对分离腔内和管路内的气泡进行彻底清除。在缓冲液注入前,超纯水应以低速不断地注入分离腔内。

(2) 缓冲液和样品的加注

开启冷却至所需温度,然后加注分离缓冲液和样品。选择合适的缓冲液流速和上样速率是自由流电泳分离的关键之一。一般样品速率与层流厚度的比值越小,样品带的扩散越小。样品带的扩散还与样品速率与缓冲液流速的比值成正比。

在加电压之前,还需对收集管进行仔细地抽气,保证每管畅通并流速均匀。

(3) 电泳

电压是自由流电泳另一个比较重要的操作参数。为了得到最佳的分离效率,必需在物料流通量和分辨率之间进行折中,这主要通过调节缓冲液流速和分离电压的大小来实现。

整个分离过程仅约几分钟,但作为制备电泳,其电泳时间通常为 1h 左右,在实际操作中根据样品量和收集液的量来决定。

电泳过程中如出现气泡,应停止电泳。

(4) 收集和检测

电泳结束后,取出收集管,并对每管进行光密度值的检测。根据光密度值,画出分离曲线。对出现的峰可进行进一步分离或分析鉴定。

电泳结束后的分离腔应以超纯水清洗,如长期不用,可用含 0.01%的叠氮钠进行防腐处理,并取出电极膜、间隔条等作适当保存。

四、分离模式

自由流电泳有 4 种最基本的分离模式，它们分别为区带电泳（zone electrophoresis，ZE）、等速电泳（isotachophoresis，ITP）、等电聚集（isoelectric focusing，IEF）和电场梯度电泳（field step electrophoresis，FSE）。可以说，这 4 种模式各有千秋，若能针对不同样品适当地选择不同的分离模式或利用一些联合型模式来进行分离纯化，往往可使整个工作收到事半功倍的效果。

（1）区带电泳（ZE）

在区带电泳中，均一的分离介质从一个输入口连续泵入分离腔并在腔内形成薄层液流，待分离样品则在其稍下游处进入分离腔，根据样品中各组分的电荷-质量比的差异进行分离，其中所使用的分离介质的 pH 和电导都是均一的。这是一种最常用、最简单方便的分离模式。在这种模式中，选择合适的分离 pH 值是整个分离成功与否的关键，介质电导通常偏向低值。

区带电泳最具吸引力的优点是其分辨率很高，回收率也高，且其设置简单，易于操作，适合于从复杂的混合物中分离纯化某一种物质，同时纯化几种物质也有报道。其缺点是产率不高，而且有条带扩散现象和稀释效应。

图 7.12 为标准蛋白质的 ZE 分离图谱。分离缓冲液和电极液均为 pH6.5，10mmol/L 三乙醇胺，含 10% 甘油。样品浓度为 1%。

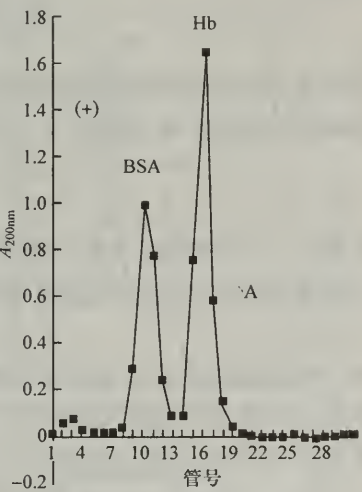


图 7.12　自由流区带电泳分离血红蛋白和 BSA

分离腔：205mm×60mm×0.8mm，缓冲液流速：2.0ml/min，样品流速：1.0ml/h，电压：300V，电流 32mA

（2）等速电泳（ITP）

在等速电泳中，被分离样品作为一条与先导离子（具有最大电泳迁移率的离子）和尾随离子（具有最小电泳迁移率的离子）溶液平行的条带同时进入分离腔，在电场的作用下，各种离子根据它们的电泳迁移率差异重新按序排队，具有较大迁移率的离子则靠近先导离子，具有较小迁移率的离子则落在近尾随离子一侧，从而达到分离。当达到稳定状态后，所有各自分开的离子都将以相同的速率移动，这样，所有离子将依它们的种类分为相应数量的彼此毗邻的条带。为收集方便，还必须加入合适的间隔物质，使各物质之间具有明显的界限。

等速电泳是代替层析的电泳对应物。其分离过程达到稳态的时间与所加的电压的大小成反比；与混合物中各组分的电泳迁移率差异的大小成反比；与样品的浓度大小则成正比。已经分开的样品中各组分的条带宽度和浓度依赖于各自迁移率的大小，而且被先导离子的浓度和迁移率大小所控制：当先导离子的浓度和迁移率均增大时，各组分的条带宽度将减小而浓度则增加。

ITP 的分辨率适中，样品带的扩散效应较小，适于分离成分不复杂的，而且其中的几个组分需同时纯化的混合物。其缺点是系统设置和操作都比较复杂，需要一定的经验来

建立合适的体系。所以,等速电泳在蛋白分离纯化方面的应用远没有其他几种工作模式那样广泛。

(3) 等电聚焦(IEF)

等电聚焦是将样品置于 pH 梯度中移动直至净电荷为零时停滞不动而聚焦,使含不同等电点的组分得到分离的过程。不同于其他 IEF 的是,在连续流等电聚焦中,pH 梯度的形成可有多种方式。常用的是直接将含有一定 pH 范围的两性电解质和分离介质泵入分离腔,在电场的作用下,形成一天然的连续的 pH 梯度,待稳定梯度(大于 30min)形成后再进样作进一步的聚焦分离,整个过程需 2h 以上。还有一种是在不同的输入口以平行的方式同时泵入不同 pH 的几种介质使形成一人工的不连续的 pH 梯度,通常仅需 5~10min 就能达到稳定梯度,样品也可与介质同时输入,但这种方法样品的聚焦程度不如前者。Weber[44]认为,利用多种不同 pH 值的缓冲液所形成的这种人工 pH 梯度具有足够强的条带聚焦效应,可以抵消部分样品带的流体变形效应,从而提高自由流电泳的分辨率。

等电聚焦过程本身是一样品聚焦和浓缩的过程,因此其条带扩散现象较小,常可用来进一步纯化样品。但其产率较低,聚焦所需时间长;蛋白质在等电点附近溶解度小,易沉淀析出而导致分离失败。另外,商品化的两性电解质混合物比较昂贵,并会与某些蛋白发生相互作用,同时,收集的样品中的两性电解质的去除也是问题,且不适合用来分离活性细胞。

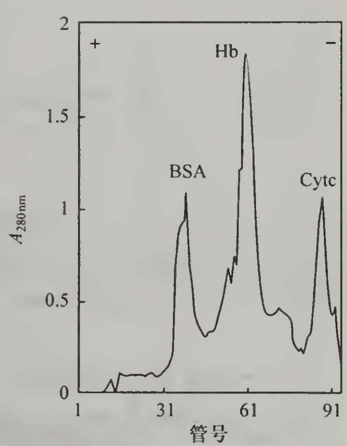


图 7.13 自由流等电聚焦分离 BSA,血红蛋白和细胞色素 c
缓冲液流速:110ml/h,样品流速:1ml/h,电压:600~1 200V,电流:15mA

图 7.13 为标准蛋白质的等电聚焦分离图谱。分离介质为 0.5% 两性电解质,含 3% 甘油;阳极电极液为 100mmol/L 磷酸,阴极电极液为 50mmol/L 氢氧化钠。样品浓度各为 1%。

(4) 电场梯度电泳(FSE)

在这种模式中,将两种缓冲组分相同但浓度相差 10 倍以上的分离介质(pH 值相同)平行泵入分离腔,通常在靠近电极处为电导高的介质,而腔中心区域为电导低的介质,样品在低电导介质区域内连续泵入分离腔,并利用腔内形成的不连续的电导梯度进行分离。当样品进入电场后,由于低电导区域的场强大大高于高电导区域的场强,欲分离物质根据其电荷迅速向相应的电极侧迁移,当迁移至电导交界处时,由于场强的骤降,迁移速率迅速变小而在界面附近浓集起来。

电场梯度电泳的一个显著特点就是其物料流通量很高,这主要是因为它所需的停留时间一般都比较短,一般小于 5min 即可。另一个特点是在不同电导区的边界处有很强的浓缩效应。在高电压、短停留时间的操作条件下,提高进样速率以及增加进样宽度都可增加物料流通量从而提高产率。尽管增加进样宽度将导致样品峰的拖尾现象,降低浓缩效应,但对产率的提高影响不大。在电场梯度电泳中虽然其浓缩效应可使样品带的扩散受到限制,但其分辨率仍显著低于区带电泳。

在这种分离模式中,分离结果的好坏与样品的处理相关。样品要求具有较低的电导值,因为样品的电导增加少许即可使浓缩效应降低,物质的迁移速率同时降低。另外,由

于其浓缩效应剧烈时会引起沉淀现象，因此对样品的起始浓度亦有一定的限制。

除上述 4 种模式外，还可将不同分离模式联合起来使用，最常见的是联合的电场梯度区带电泳(combined field step-zone electrophoresis, FSE/ZE)。在这种联合模式中，介质的设置和分离原理均与电场梯度电泳类似，惟一不同的是以区带电泳的细带进样，而不是电场梯度电泳的宽带进样。这种模式的分辨率比电场梯度电泳高，同时其条带扩散现象比区带电泳有明显改善，且其产率适中，因此适合于分离成分不太复杂的混合物。

图 7.14 为天花粉蛋白的最终纯化图谱。分离体系 pH 为 7.9，分离缓冲液为 10mmol/L 三乙醇胺，含 5%甘油；电极液为 100mmol/L 三乙醇胺；稳定介质为 100mmol/L 三乙醇胺，含 5% 甘油；反向介质为 5%甘油。样品蛋白浓度为 4.2mg/ml。

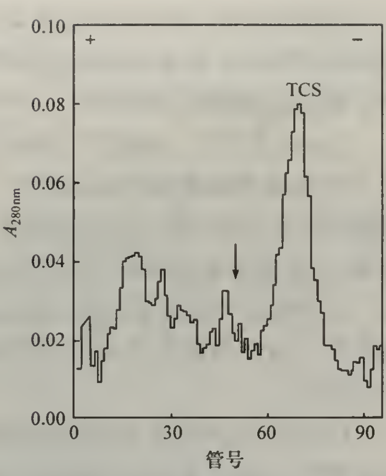


图 7.14　自由流电泳 FSE/ZE 模式纯化天花粉蛋白

缓冲液流速：500ml/h，样品流速：2.3ml/h，电压：1 600V，电流：75mA

五、自由流电泳的应用

自由流电泳由于其连续和温和的分离条件，适合于多种生物活性物质的分离，从小分子、蛋白质、DNA、脂质体、膜物质、细胞器到完整细胞等，涉及到生物学、生物化学、细胞学等多个学科领域。

(1) 小分子

生物样品中的一些大分子物质可通过透析、超离心、凝胶色谱等方法除去，但一些小分子的去除反而比较困难，而自由流电泳则为其分离提供了可能，并且经过 CFFE 分离的样品可直接用于红外、核磁共振和质谱的分析。

ITP 模式适合于分离小分子。Hirokawa 用 α-羟丁酸或酒石酸为介质从核燃料废弃物中以 ITP 模式分离铂分子，之后以颗粒诱导的 X 射线方法测得其中主要的阳离子的回收率为 100%；还以醋酸分离体系分离带负电的苋莱红染料和荧光染料。FSE 模式成功地应用于经离子对色谱分离后组分中混有的离子对试剂的分离。

(2) 多肽和蛋白质

不管在地面还是空间，蛋白质、酶或一些酶解产物已成为 CFFE 广泛应用的对象，尤其是蛋白质，还经常被作为模型样品来研究 CFFE 所包含的现象以及现象的消除，并用来试验分离装置的可行性。

Hoffstetter-Kuhn 和 Wagner[45]把从酵母蛋白粗提物中纯化乙醇脱氢酶(ADH)作为模型，深入地研究了放大型 CFFE 电泳在工业领域进行酶的纯化方面的潜力和限制因素，指出 FSE/ZE 在蛋白质纯化上的潜力。他们对介质的 pH、电导、停留时间、分离电压以及样品进样速率、浓度等进行了探讨，经条件优化后，以三个进样口同时进样，可得到 2.7 g/h的物料流通量，样品提纯 4.7 倍，总产率达到 90%以上的分离结果。

Barth 等[46]人对 t-PA 进行了分离纯化，使用 5mmol/L 的醋酸-醋酸铵介质，同时在分离介质中加入 0.05%的吐温 80 以减小吸附现象，得到 80%以上纯度的样品，并可直接用于糖基化研究和 N 端测序。

Loseva 等[47]人，纯化了几种蛋白质和多肽：蜜蜂毒液、肿瘤坏死因子、β-白细胞介素、γ 干扰素和超氧化物歧化酶等。应用 CFFE 分离蛋白质的例子还很多，如鸡和鹅的红细胞中组蛋白的分离、HeLa 细胞中 NADH 氧化酶的抗癌性磺酰尿抑制剂和抗性形式的分离、扁豆凝集素的制备、干扰素的纯化、天花粉蛋白的纯化[39]等。

空间电泳分离的蛋白质，目前以模式蛋白质为多，如血红蛋白、血清蛋白、白蛋白、细胞色素 c、血红细胞生成素、胰蛋白酶抑制剂等，也有分离培养基质中的目标蛋白的。

(3) DNA

DNA 片段的电泳分离已很成熟，主要是凝胶电泳。以 CFFE 分离 DNA 并不多，但自 Bier 等[48]人将几种成纤细胞株中分离的中期染色体以 CFFE 研究其电泳迁移率，并以细胞遗传学方法和流式细胞术都证明了 CFFE 分离对染色体的形态和稳定性均无影响后，人们开始注意 DNA 的 CFFE 分离。最近，日本在空间成功地分离了 *C. elegans* DNA[49]，取得令人兴奋的结果。

(4) 膜物质、生物颗粒、细胞

几乎没有一种更合适的分离方法可替代 CFFE 来分离膜物质、生物颗粒和细胞。不同于其他分离对象，为了保持样品的生物活性，在分离过程中的 pH 必须保持在其生理容忍范围内，同时为了提供等渗环境，有时还需在介质中加入一定的蔗糖、葡萄糖等。

膜物质的 CFFE 分离引起了很多生物化学家和生物学家的广泛注意，许多膜物质已用 CFFE 得到分离，其中人们研究较多的有内体、高尔基体、质膜、液泡、线粒体和溶菌酶以及一些特殊的膜物质如披网格蛋白小泡、胰岛素响应的葡萄糖转运活性小泡等。

以 ITP 模式可对生物颗粒和细胞器进行分离。Volkl[50]将小鼠肝内细胞器被分离成过氧化物酶蛋白质和微粒体的蛋白包含体；Lackner 等将胆固醇、甘油三酯、卵磷酯、神经鞘磷脂和转脂蛋白等主要颗粒群体进行了分离纯化及功能研究；Spaans 等[51]人纯化了猪胰酶原颗粒。还有人从云杉中分离出完整的叶绿体等。

Hymer 等[52]应用 CFFE 在微重力下从培养的小鼠脑垂体前体中分离纯化了含有生长激素的胞质分泌颗粒，并与地面结果相比，发现在轨细胞有更高的电泳迁移率，表明暴露于微重力下的细胞表面静电荷密度会发生改变。

对于细胞的分离由于其所要求的生理环境很苛刻及不易保持其生物活性，使一些常规的分离方法显得无能为力，然而不同的细胞其表面积的负电荷密度一般都有差异，因此可用 CFFE 来进行分离，同时因 CFFE 的独特性，使其备受细胞学家和生化学家们的青睐。人们已开展地基和空间的细胞分离研究工作，并有数百种细胞应用 CFFE 得到分离。

早在 60 年代，Hannig 和 Zeiller 等人就已开始进行免疫系统中各种细胞的分离工作，在分离分化和激活不同时期的 B、T 淋巴细胞的过程中证明了 CFFE 是一种强有力的研究方法。Bauer 和 Weber 在细胞生存力和电泳分离条件的关系上作了大量的研究，并提出介质中氯化钠浓度至少达到 35mmol 以上，才可能满足分离后细胞的生存力的要求；同时，他们还发展了三种很好的细胞分离缓冲液系统[53]。现在，B 和 T 淋巴细胞已经可以通过自由流电泳大量地进行制备。

精子是一种特殊的细胞,它本身就具有移动能力。但由于带有 X 染色体的精子与带有 Y 染色体的精子具有不同的表面电荷密度,所以同样可用自由流电泳对其进行分离。Kaneko 等[54]人和 Engelmann 等[55]人用自由流电泳均得到了 100%纯化的含 X 染色体和 80%纯化的含 Y 染色体的人精子。

自由流电泳还可用于临床上的血液细胞快速检测,以监测身体的病变情况。Valet 等[56]人通过对绵羊红细胞群体动力学的研究发现,在幼羊的血液中存在三种红细胞群体,其电泳图谱表现为在幼年动物中有两个峰,而在成年动物中只有一种红细胞。将每天的电泳图谱依次排列起来,可清楚地看到整个变化过程。

图 7.15 为我们实验室对人和兔红细胞的分离图。分离缓冲液为 15mmol/L 三乙醇胺,4mmol/L 醋酸钾,10mmol/L 葡萄糖,240mmol/L 甘氨酸,30mmol/L 蔗糖,pH7.2。

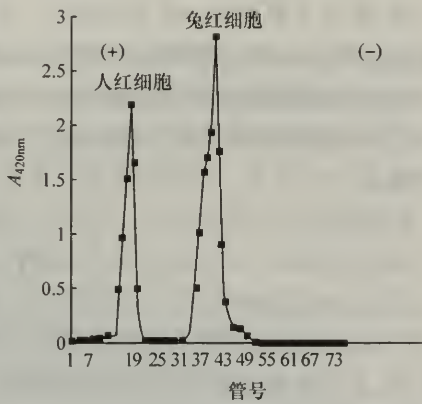


图 7.15 自由流电泳分离人和兔红细胞
缓冲液流速:2.7ml/min,样品流速:1.38ml/h,
电压:400V,电流:63mA

为了能对那些电荷密度差异不大的细胞进行分离,人们还发展了抗原特异性电泳细胞分离法和以后的免疫自由流电泳,使细胞的电泳分离更广泛。

在空间进行过的细胞电泳对象,以不同种类的红细胞为多,也有肾细胞、脾细胞、淋巴细胞和骨髓细胞。

六、自由流电泳在微重力下的特点

颗粒沉降、小滴沉降和热对流是影响自由流电泳的三种重力现象,在微重力条件下这些现象几乎完全消失,这使 CFFE 的放大在空间有可能得到实现,而且也只有空间条件才是消除重力依赖的影响因素的惟一机会。

在地基条件下,为了减小热对流,分离腔的厚度必须保持在 0.5~0.8mm 内,并且还必需有十分有效的冷却。另外,样品流的直径约为分离腔厚度的 1/3 时,样品在介质流的 Posseuil 流影响下可达到最大速度迁移。上述所有条件都限制了样品的上样量。微重力条件下,分离腔内的热对流几乎完全消失,因此,从理论上讲可大大增加分离腔的间隙,以提高样品上样量。然而,仍有两个对立的因素限制了 CFFE 在空间的无限放大,使放大终究有一定的限度。因为:其一,在冷却玻璃板和分离腔一半厚之间仍存在由焦耳热引起的温度梯度,所以在空间条件下也应控制分离介质的温度,使其在蛋白质等活性材料的变性温度以下。其二,冷却能力限制了分离腔厚度的上限。通常,腔厚度增加将明显降低腔内液流的散热效果,加大腔内温度变化;而温度增加 1℃,会使电泳迁移率增加 3%,从而降低了过程的分辨率。考虑到这些,介质流的厚度实际上最大为 3~5mm。然而,即使这样,在分辨率保持基本不变的情况下,物料流通量仍能增加约一个数量级,甚至更高。

与地基条件相比,在微重力下由于沉降现象的消失可使用更高浓度的样品上样,使物

料流通量进一步提高 5～10 倍(依赖于被分离的生物材料)。这种提高可以认为是由于在微重力下避免了微滴的形成而实现的。

许多学者以理论模型预言在微重力下将能使用更大的分离腔、更高的场强、得到更好的分辨率以及更大的样品流通量。在过去 20 年的空间电泳中,对蛋白质和生物大分子的分离,证实了微重力条件主要提高了流通量和分辨率,如美国的 CFES 实验指出空间流通量将提高 718 倍;对颗粒的空间分离证明了运用高场强和大分离腔的可能性,同时也证实了对流和沉降消失后对分离有利的理论分析。此外,在空间对细胞的分离也证实了在空间可有更宽的操作条件;这些都至少证明了微重力的确为自由流电泳提供了相当优越的环境。

七、展　望

自由流电泳的发展经历了较长一段时间,逐渐向人们展示了它的潜力,一些其他分离方法,尤其在制备级上无法比拟的潜力。同时,它还是一个合适的开展电泳理论研究的对象,如探讨热对流、各种动力学变形、微重力等对电泳的影响。而空间电泳在分辨率和流通量上的提高,更吸引了众多的生化学家的注意力。但由于空间电泳的高成本,决定了候选样品是那些在地面不可能在纯度和量上达到要求的样品,或是那些具有极高商业价值的样品,如其年产量价值至少达到 1000 万美元以上的目标产品。

在未来空间站上,生物样品的纯化不只是一个简单的过程,它可组合不同的生物分离技术通过一系列的生物过程最终得到设计的目标产物。在此之中,细胞电泳可用于分离空间电融合所产生的杂交瘤细胞或其他细胞,纯化的细胞再作为空间生物反应器中的种子细胞,在培养基质条件下,使之产生大量细胞,然后维持细胞在有利于产生分泌激素、抗体等产品的条件下使之富集,最后将富集的目标产品以空间电泳一步纯化。在此方案中,将这些生物技术组合在一起对空间筛选新样品是相当有潜力的。微重力研究也早已证明了其中每一过程的优越性,其中空间电泳、空间电融合和空间生物反应器已在空间分别得到飞试,而流式细胞仪也已为空间站的使用在不断发展。

自由流电泳用于蛋白质组研究已有报道[57]。蛋白质组的研究实质上是在细胞水平上对蛋白质进行分离和分析,从技术方法上,目前主要以双向凝胶电泳分离和质谱分析为支撑技术(见有关章节)。传统的双向凝胶电泳由第一向等电聚焦和第二向 SDS-PAGE 组成,虽然其一次能分离到上千个蛋白质点,但有其固有的缺陷,如严格限制样品的上样量(通常 50μg～1mg),第一向到第二向存在低转移效率,样品在样品缓冲液中的低溶解度,许多蛋白尤其是膜相关蛋白和低拷贝数的蛋白得不到检测等等。因此 Simpson 等[57]人以自由流电泳的等电聚焦模式结合 SDS-PAGE 对结肠癌细胞系 LIM 1215 进行了分离和质谱分析,鉴定到 36 个蛋白点。此法一方面可提高样品的上样量(连续进样),另一方面可提高样品的回收率,自由流电泳收集的样品仅有 1%用于 SDS-PAGE,剩余的 99%可用来研究目标蛋白质的性质及其进一步的结构特性,研究细胞信号传导的复杂性,作进一步的细胞作图的蛋白质组学研究。另外,以自由流电泳取代传统的凝胶等电聚焦进行的二向电泳,还可用来分离细胞器。我们实验室利用自由流电泳的区带模式结合混和酶解及质谱分析,也得到一定的结果。

我们相信,随着电泳仪和操作条件的不断完善,以及将来空间站所提供的便利,自由流电泳必将发挥越来越大的作用。

第三节 液相等电聚焦

蛋白质的分离方法可依赖于蛋白质的各种不同性质来区别,如质量、形状、电荷,疏水性、亲和性等,或者结合这些因素中的几项。大多数工作者在聚丙烯酰胺凝胶或琼脂糖凝胶上电泳,其目的是分析,如等电聚焦、SDS-PAGE、双向电泳,都有很好的分辨率,广泛用于生物学研究和质量控制中。而用这些技术进行制备量是非常困难的,因为上样量通常为 0.1～2μg,而且很难从凝胶上回收蛋白质。与分析电泳相比,制备电泳的分辨率较差,还存在不少困难,譬如由电场产生的焦耳热,是正比于电泳体积;尽管可以采取冷却的方法,但冷却的效果又反比于其表面积,往往难以抵消产生的热量。等电聚焦作为纯化蛋白质的方法,不仅具有高分辨率、样品高浓缩效果,以及可以获得高的回收率,而且可以不改变样品的物理化学性质和生物活性。因此成为蛋白质分离纯化的手段之一。

目前,用于蛋白质制备目的的等电聚焦有多种,根据 IEF 时所用的支持介质的种类不同,基本上可以分两类:一类是在液体介质中的制备等电聚焦,是本节的主要内容;另一类是在凝胶介质中的制备等电聚焦,这里不作介绍。

制备型液相等电聚焦的特点是无载体。采用无载体系统,仅在液相中进行等电聚焦,就不存在从载体上回收样品的问题。此外,上样量可从几微克到克级,满足了制备级的要求。在无载体情况下,要解决样品溶液和分离液体介质的扩散和电泳时产生热量而引起的热对流。否则,犹如牛奶滴到水中,样品反而稀释,更谈不上等电聚焦纯化样品了。

我们实验室应用了旋转式等电聚焦(rotofor)、循环式等电聚焦(recycling isoelectrofocusing,RIEF)、连续液相等电聚焦(continuous IEF,CIEF)电泳技术。本节将几种类型的液相等电聚焦作一简单介绍。

一、旋转式等电聚焦

(一) 概　况

1983 年 Bier 报道一种水平式聚焦柱,它由 20～30 个平行而有连续的部分组成,电极在两端,它们之间管道是相通的。整个聚焦室周围有一陶瓷冷却管,聚焦室围着中央冷却棒旋转(1r/min)[58],阻止了热对流和地心引力对蛋白质分子的影响,这种想法最早在 1967 年由 Hjerten 提出:重力影响可以用电泳柱围着它的水平轴旋转而减小。

Rotofor 装置是 Bio-Rad 公司生产的制备 IEF 仪,是根据区带对流等电聚焦仪改进的,是一种制备级的无载体等电聚焦仪,为蛋白质的纯化提供了新的选择对象。其优点是在反复的聚焦过程中,使蛋白质纯度增加。举例来说,如果某种有意义的蛋白质经第一次分离后(两性电介质 pH3～9),发现分布在二个或三个收集管内,可对此样品采用相应的更窄 pH 的两性电介质(pH4～6)进行第二次聚焦,使纯度增加。该仪器的特点如下所述。

①等电聚焦可直接在溶液中进行,不需加稳定介质。聚焦室含有 19 个平行的聚酯膜隔板,使聚焦室分隔成 20 个不连续的间隔部分。等电聚焦过程中可控制蛋白质的扩散。

②在聚焦室的中心有一指形的陶瓷冷却管，它与聚酯膜隔板的膜芯相联，可阻止电泳过程中的热对流。

③在 IEF 过程中聚焦室可以转动，而使地心引力被消除。

④等电聚焦结束后，可将含有 20 个针管的收集装置与聚焦室内 20 个间隔室相连，通过真空吸引能在几秒钟内收集到 20 个间隔室内已通过 IEF 分开的蛋白质而不被混合，保持 pH 梯度和等电点蛋白质。

由于以上一些特点使 Rotofor IEF 制备电泳仪分离蛋白质获高分辨率成为可能[59]。

（二）Rotofor 制备电泳仪操作步骤

(1) 组装

根据 Rotofor 制备电泳使用说明书组装好仪器(图 7.16)。

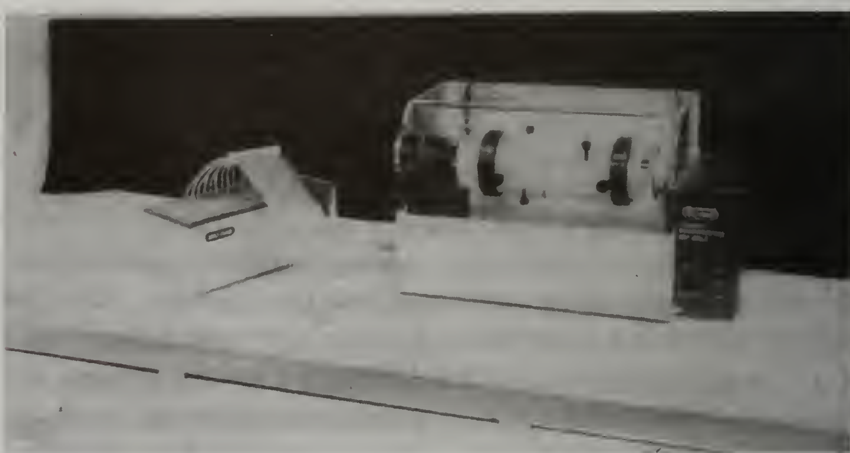


图 7.16　Rotofor 制备电泳仪和取样装置

(2) 加样

加样品于聚焦室内(注意不要带入气泡；样品最小体积 30ml，最大体积 55ml)，蛋白质样品制备：根据蛋白质的性质可用以下几种溶液。

溶液 1：样品溶于 2% 两性电解质中(pH 范围根据需要选用)。

溶液 2：样品溶于 3mol/L 脲和 2% 两性电解质中(pH 范围根据需要选用)。

溶液 3：样品溶于 9mol/L 脲，2.5ml β-巯基乙醇，10mmol/L NaCl，2% 两性电解质(pH3～10)，1% 两性电解质(pH5～7)(pH 根据需要选用)。

每次 IEF 负载量大约 1.5g 蛋白质。脲和非离子去垢剂可以保证蛋白质的溶解。

(3) 电泳条件

负极电极液　0.1mol/L NaOH；正极电极液　0.1mol/L H_3PO_4。

样品加入聚焦室后，盖上盖子。聚焦室旋转 4℃ 预冷 10～15min 后，加稳定电压至 300V，检查功率。最后以 12W 恒定功率进行电泳。电泳过程中，电压会不断上升(最大电压 2 000V)，电流不断下降并恒定，4h 完成电聚焦。

注意：如果起始电压过高可在样品液中增加 NaCl 浓度，使电压降低。功率最高只能保持 12W(4℃)，过热会损坏聚焦仪。

(4) 取样

电泳结束后，停止仪器的转动；聚焦室带有取样孔的一面朝下，将取样装置连接到减压水泵，盖上取样室盖子，盖上的 20 支针头同时刺破聚焦室取样孔的膜而伸入聚焦室内；通过真空吸引可在几秒钟内收集 20 个样品于取样室内的 20 支试管中。

(5) 电泳检查电聚焦样品

20 个 IEF 后的样品可直接测定每一管的 pH。取小量可分别进行薄层 IEF(或盘状 IEF 和 SDS-PAGE 电泳检查蛋白质分子质量的检查)。如果目的蛋白质在某一管或分散在几管中样品用水稀释后可再次进行 Rotofor IEF 一次，直至纯的蛋白质的获得。

(6) 等电聚焦后的载体两性电解质的去除

①由于两性电解质的分子质量较小，因此透析的方法可以去除它们。

②蛋白质装于透析袋内，通过电泳，使两性电解质泳出透析袋。

③如等电点沉淀出的蛋白质可用离心法收集，用等电点溶液洗去电解质。

④用 60%～100%硫酸铵，在对蛋白质稳定的 pH 条件下沉淀出蛋白质，然后用同浓度的硫酸铵再洗 3 次，去除两性电解质。

⑤可用 Sephedex G-25 凝胶过滤法，去除两性电解质。

⑥超滤技术的应用。

我们研制类似的简化仪器用于分离蛋白质，基本结构和 Rotofor 相似，但聚焦柱由 12 个动片和 12 个定片间隔组成。在电聚焦结束后，动片转动一个角度，便将原来相通的聚焦池分成互不相通的 12 个独立的小室，然后分别取出样品。图 7.17 是我们用自制的仪器制备卵白蛋白和细胞色素 c，在收集后再用 Mini-IEF 鉴定[60]。

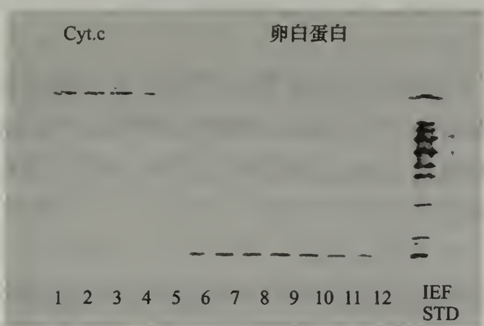


图 7.17　卵白蛋白和细胞色素 c 的 IEF 分离

样品各 10mg，两性电解质 pH3～10，

电压 230V，持续 24h(无冷却系统)

二、循环式等电聚焦

1. 原理及其组成

后来 Bier[58]又设计了一种循环式等电聚焦，它的原理图如图 7.18 所示。

以 RF3 为代表的循环式等电聚焦仪主要由 5 部分组成。

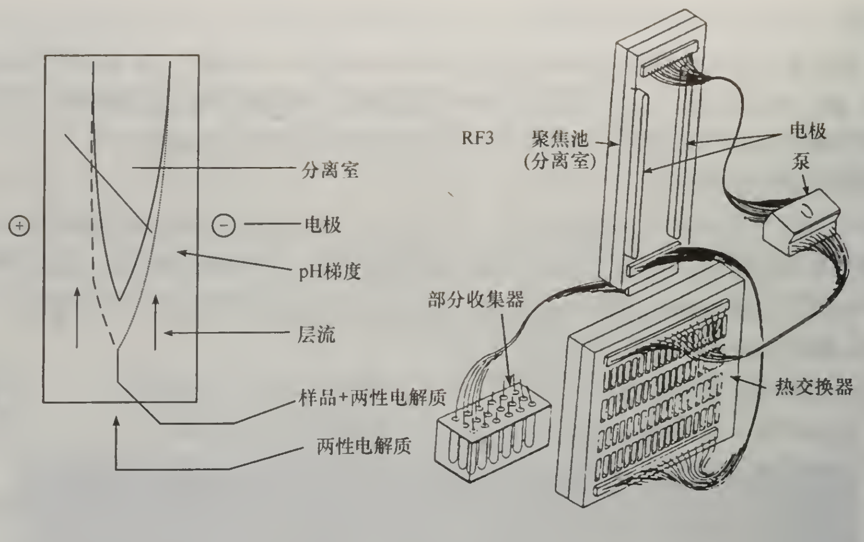


图 7.18　RF3 原理图和装置

(1) 聚焦分离室

聚焦室是循环自由流等电聚焦仪的核心,左右两侧分别为正极和负极小室,离子交换膜既使其与中间的聚焦室分开,又通过膜仍可相通,这样电极上产生的气泡就不会干扰中央聚焦室已分离好的介面。电极室的酸碱也以高速循环流动,可将气泡赶走而不致影响电场。中央分离室循环流动的两性电解质由下而上,在正负电极所提供的电场中形成从正极到负极逐渐升高的近线性的 pH 梯度,具有不同等电点(pI)的蛋白质"对号入座",从而各自分离。

聚焦室是矩形的,两平行的板之间厚度 0.75mm,液体在其间为一极薄的幕状流,在蠕动泵作用下以较快速度从底部 30 孔进入聚焦室,然后由顶部相应的 30 孔流出,再经外部管道流过冷却系统得到降温,如此反复循环直至聚焦完成,此时转动三相阀门,分离聚焦好的样品则进入收集器。

(2) 热交换器

等电聚焦过程中不断产生和积累的焦耳热是影响电泳效果的不利因素之一,大电流影响尤为严重,同时产生的焦耳热正比于装置的体积,而冷却效果只正比于它的表面积。循环自由流等电聚焦引进外部冷却系统,聚焦过程可以控制在合适的温度下进行,避免了温度过高而可能对样品生物活性造成不利影响,有效保证了聚焦后样品生物活性的回收。控制温度的作用还在于显著地减小对流,防止扩散,使聚焦达到最佳效果。

(3) 蠕动泵

蠕动泵是液体循环流动的动力来源,有两种泵,一种为双向可调速,用于驱动两性电解质载体和样品的循环流动,改变泵的转动方向和速度可以进行进样、驱赶系统中气泡、收集样品等目的。另外一组泵驱动聚焦池两侧电极室中电解液的循环流动,电解过程中可能产生的气泡在循环流动过程中及时得到排放。

(4) 排气泡装置

排气泡装置在循环流动过程中的主要作用是排除系统在聚焦过程中可能产生的气

泡,还可以在样品体积较小时在此进行加样。

(5) 收集器

聚焦后的两性电解质 pH 从第 1 管到第 30 管呈线性分布,样品则按其等电点分配于相应的管道中循环流动,循环自由流等电聚焦采用 30 只收集管组成的收集器,在蠕动泵的作用下分别对应收集,避免了聚焦后目标样品的再混合。

2. RF3 的操作步骤

①将配制好的阴极(0.1mol/L 氢氧化钠)和阳极(0.1mol/L 磷酸)电极液加注到电极储杯内,将两性电解质溶液(1%两性电解质,10%甘油,105ml)以针筒注入分离腔内,并注意将腔内的气泡敲击赶出。

②预聚焦:将电参数设置到最大(150W,400mA,1 500V)进行预聚焦 30～45min。

③加样:待电流下降并稳定于某一值后,以针筒注入欲分离样品(5～10ml)。

④聚焦:调低电压至 1 000V,聚焦 1h;之后 500V,聚焦 0.5h。

⑤取样:聚焦结束,同时关泵和高压,并将聚焦好的溶液赶至 30 管收集管内。

⑥检测:收集的样品作 pH 值和光密度值测定,并可做进一步的分离,如将每一管继续做 SDS-PAGE 电泳分离,或将其中的一管或几管重新进行不同 pH 范围的等电聚焦以获得更纯的样品。

⑦仪器的清洗:聚焦结束后的分离腔应及时进行清洗,保证以后的正常使用。依次以 500～1 000ml 水,300～400ml 0.1mol/L 氢氧化钠,1 000～1 500ml 水,200ml 10%酒精冲洗分离腔,排空电极槽并以水冲洗,最后放松泵。

3. RIEF 的应用

①用于酶的精制纯化[61],从米曲霉中初步提纯的 DNA 核酸酶,经 RIEF 分离,酶活性集中在第 14 管,活性回收 82%。

②除去干扰素中的热源[62],α-干扰素-2 制剂中含有热源,影响其临床应用,Nagabhushan 等人先用 pH3～10 的两性电解质分离,发现样品集中的 5～8 管中有抗病毒活性,但仍含有一定量的热源,根据这几管的 pH 在 5～7 之间,采用 pH4～7 的两性电解质第二次聚焦,可将干扰素和热源分开。

我们实验室以标准蛋白进行 RIEF 的分离,结果见图 7.19。

分离样品:1%细胞色素 c、1%Hb、1%BSA。

(样品需经色谱柱脱盐或含低盐浓度,用 0.45μm 微孔滤膜过滤)

电极缓冲液:负极:0.1mol/L 氢氧化钠　正极:0.1mol/L 磷酸。

分离室缓冲液及电泳条件:配成 1%两性电解质(军事医学科学院产品),含 10%甘油,共 105ml 注入分离腔,进行预聚焦 30～45min,电压 1 000V。然后加样再进行聚焦,每隔 10～15min 记录电流和功率,直至电流衡定,调节电压 1 200V,继续聚焦 10～15min,最后收集。

样品量大于 10ml,可不进行预聚焦,加去离子水至 102.5ml(含 10%甘油),再加 40%两性电解质 2.5ml,直接进行聚焦。

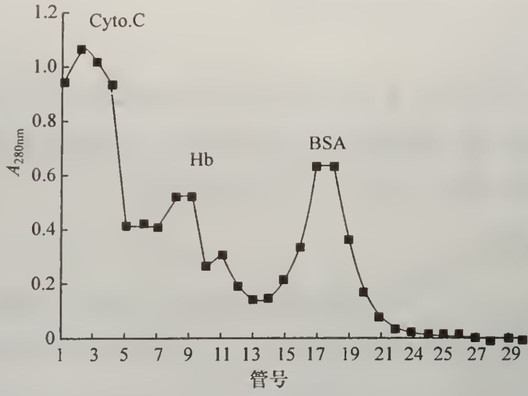


图 7.19　BSA、血红蛋白，细胞色素 c 的分离纯化

4. 讨论

等电聚焦早期用单个氨基酸时，pH 范围窄，缓冲能力差。目前大多用两性电解质，这是一种多氨基多羧酸，具有趋于连续改变的氨基羧基比，pH 值连续，能得到较平滑的 pH 梯度。

两性电解质对某些酶活性有影响[63]，限止了它的应用。因此也有用几种氨基酸衍生物代替两性电解质建立阶梯式 pH 梯度的探讨[64]，Righetti 则采用一种膜建立固定化的梯度(immobilized gradients)以代替液相两性电解质[65]，最近 Hoefer 公司已有此类商品化的仪器供应。

两性电解质的质量对结果有影响，有些国产两性电解质的紫外吸收很高，干扰了蛋白质的定量，因为在聚焦后观察蛋白质的分布，以测 280nm 的光吸收最方便，又不损失样品。如两性电解质的紫外吸收很高，以致无法测到 280nm 的光吸收。有些国产的两性电解质可能含有盐，因为在同样实验条件下，电流明显偏大，这也有影响。

关于如何去除 IEF 纯化的蛋白质中的 Ampholine，Nilsson 等[66]人报道，用 ^{14}C-Ampholine 实验，用 Sephadex 分 Ampholine 和蛋白质有两个弊端，一是回收低，二是样品在此过程中被稀释，作者采用硫酸铵沉淀法，原来 0.5mg/ml 的 α-Haemolysin 和 5mg/ml BSA，Lysozyme 中放射性为 $(2\sim5)\times10^6$ cpm/ml，在重复沉淀 3 次后，再把蛋白质溶解在 0.1mol/L 磷酸缓冲液(pH7.0)中，完全测不到放射性。

三、制备型连续流等电聚焦

Weber 设计了连续自由流等电聚焦仪，并使之商品化。结构和连续流电泳相同，模式如图 7.20。

0.5%两性电解质由底部连续进入分离室(1 道，最多 5 道)，其设计可保证形成稳定的平流。流经分离室时在电场作用下形成 pH 梯度。

样品由进样口连续进入，在 pH 梯度下，蛋白质根据不同等电点(pI)聚集，分离过程于 5～10min 即可达到平衡。由于设计是连续性的，在达平衡后连续进样，连续收集，因此

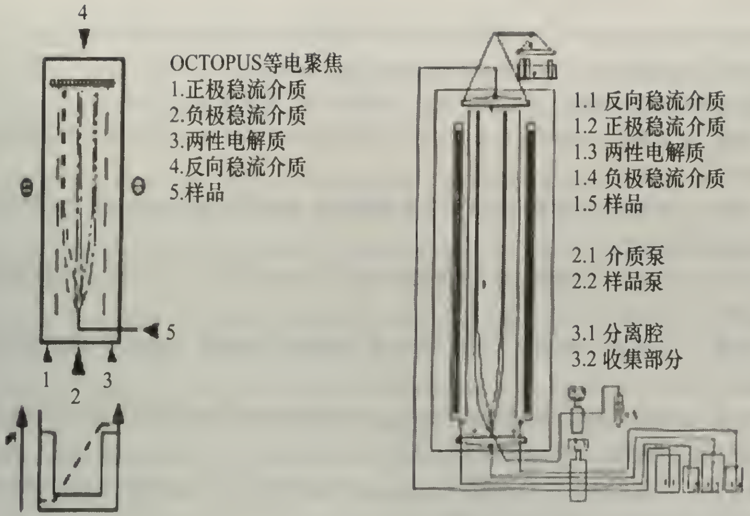


图 7.20 连续流等电聚焦原理图和仪器结构图

从理论上讲,制备量应是很可观的。

Weber 设计的仪器长 500mm,宽度 100mm,厚度 0.5mm,在分离室的后面有冷却系统,在 100mm 宽度上有 96 收集管。

我们以标准蛋白质进行了 CIEF 分离,结果见图 7.13(本章第二节)。

(宋金芳 屠红旻 龚蓁蓁)

参考文献

1 邵晓霞, 夏其昌. 第二代液相电泳. 生命的化学, 1995, 15(6):36~39

2 Jorgenson JW. Electrophoresis. Anal Chem, 1986, 556:743A~758A

3 Grossman PD. A practical introduction to free solution capillary electrophoresis of proteins and peptides. Am Biol Lab, 1990. 35~43

4 Hjerten S. Zone broadening in electrophoresis with special reference to high performance electrophoresis in capillary electrophoresis: an interplay between theory and practice. Electrophoresis, 1990, 11:665~690

5 Grossman PD, Colburn JC, Lauer HH *et al*. Application of free-solution capillary electrophoresis to the analytical scale separation of proteins and peptides, Anal Biochem, 1989, 61:1186~1194

6 Lander JP. Capillary electrophoresis:pioneering new approaches for biomolecular analysis. TIBS, 1993, 18:409~414

7 Rarger BL. Capillary electrophoresis. Curr Opin Biotechnol, 1992, 3:59~64

8 Luckey JA, Drossman H, Kostichka AJ *et al*. High speed DNA sequencing by capillary electrophoresis. Nucleic Acid Res, 1990, 18:4417~4421

9 Karger BL,Chu YH, Foret F. Capillary electrophoresis of proteins and nucleic acids. Annu Rev Biophys Biomol Struct, 1995. 24:579~610

10 Liu J. Ultrasensitive fluorometric detection of carbohydrates as derivative in mixture separated by capillary electrophoresis. Proc Natl Acad Sci USA, 1991, 88:2302~2306

11 Liu J. Design of 3-(4-carboxybenzoyl-2-quinoline-carboxaldehydeas a reagent for ultrasensitive determination of primary amines by capillary electrophoresis using laser fluorescence detection. Anal Chem, 1991, 63:408~412

12 Xia QC. The use of capillary electrophoresis for analysis of proteins,peptides and amino acids. In:DU Y C ed. Peptide

Beijing:Science Press, 1991. 243～245

13 夏其昌, 用毛细管电泳在飞克分子水平定量分析 PTH-氨基酸. 生物化学与生物物理进展, 1991, 18:303～307

14 屠红旻,夏其昌, 毛细管电泳——蛋白质,多肽分离分析的一项新技术.化学通报, 1991, (12)15～19

15 万谦, 毛细管电泳测定蛋白激酶 A 活力的新方法. 生物化学与生物物理学报,1995, 27:579～584

16 车发云, 邵晓霞, 徐来根等. 毛细管电泳非放射分析蛋白质和多肽磷酸化. 生物化学与生物物理学报, 1997, 29(4):363～370

17 车发云, 文正宇, 王克夷等. 毛细管电泳分析常见单糖的 2-氨基丫啶酮衍生物. 生物化学与生物物理学报, 1998, 30(3):251～256

18 车发云, 刘正宇, 王克夷等. 毛细管电泳分离寡糖衍生物及其电泳行为研究. 生物化学与生物物理学报, 1998, 30(5):494～500

19 车发云, 邵晓霞, 王克夷等. 毛细管电泳分析飞摩尔水平糖蛋白中唾液酸. 生物化学与生物物理进展, 1999, 26(1):84～86

20 车发云, 王克夷, 夏其昌. 寡糖 8-氨基萘基-1,3,6-三磺酸衍生物在毛细管区带电泳中迁移时间相互关系. 中国生物化学与分子生物学报, 1999,15(3):126～129

21 Che FY, Song JF, Wang KY *et al*. Comparative study on the distribution of ovalbumin glycoforms by capillary electrophoresis. J Chromatogr A, 1999, 849(2):599～608

22 Che FY, Shao XX, Wang KY *et al*. Characterization of derivatization of sialic acid with 2-aminoacridone and determination of sialic acid content in glycoproteins by capillary electrophoresis and high performance liquid chromatography-ion trap mass spectrometry. Electrophoresis, 1999, 20(14):2930～2937

23 Che FY, Song JF, Zeng R *et al*. Analysis of 8-aminonaphthalene-1,3,6-trisulfonate-derivatized oligosaccharides by capillary electrophoresis-electrospray ionization quadrupole ion trap mass spectrometry. J Chromatogr A, 1999, 858(2):229～238

24 Liu T, Shao XX, Zeng R *et al*. Analysis of recombinant and modified proteins by capillary zone electrophoresis coupled with electrospray ionization tandem mass spectrometry. J Chromatogr A, 1999, 855(2):695～707

25 刘韬, 曾嵘, 邵晓霞等. 毛细管区带电泳-串联质谱联用法鉴定多肽和蛋白质. 生物化学与生物物理学报, 1999, 31:425～435

26 Huang XL, Capillary array electrophoresis using laser-excited confocal fluorescence detection. Anal Biochem, 1992, 64:967～972

27 屠红旻, 夏其昌. 毛细管电泳及其在生命科学中的应用. 生命的化学, 1994, 14(6):36～39

28 Chen SM, Wiktorowicz JE. Isoelectric focusing by free solution capillary electrophoresis. Anal Biochem, 1992, 206:84～90

29 Roman MC, Brown PR. Free-flow electrophoresis as a preparative separation technique, Anal Chem, 1994, 66(2):86A～94A

30 夏其昌, 宋金芳. 连续自由流电泳,见:夏其昌主编. 蛋白质化学研究技术与进展. 北京:科学出版社, 1997. 151～158

31 Hannig K, New aspects in preparative and analytical continuous free-flow electrophoresis, Electrophoresis, 1982, 3:235～243

32 Bauer J. Cell electrophoresis. Boca Raton:CRC Press, 1994

33 Krivankova L, Bocek P. Continuous free flow electrophoresis. Electrophoresis, 1998, 19:1064～1074

34 邵晓霞, 宋金芳, 万谦等. 连续自由流电泳的研制和应用于分离蛋白质和细胞. 生物工程学报, 1997, 13(2):218～220

35 宋金芳, 夏其昌. 连续自由流电泳分离蛋白质的研究. 空间科学学报, 1999, 19:162～169

36 宋金芳, 李丽, 郭红艳. A3 型电泳仪研制的一些实验依据. 空间科学学报, 1999, 19:170～177

37 刘韬, 宋金芳, 夏其昌. 自由流电泳的原理和应用. 空间科学学报, 1999, 19:33～42

38 宋金芳, 夏其昌, 苏维埃. 空间电泳. 见:江丕栋主编. 空间生物学. 青岛:青岛出版社,2000. 411～443

39 Song JF, Liu T, Shen X *et al*. Application of free-flow electrophoresis to the purification of trichosanthin from a crude

product of acetone fractional precipitation. Electrophoresis, 1998, 19:1097～1103

40 刘韬，沈欣，吴高德等. 生物化学和生物物理学报，1999，31:590～593

41 Hannig K, Wirth H, Meyer BH *et al*. Free-flow electrophoresis I: theoretical and experimental investigations of the influence of mechanical and electrokinetic variables on the efficiency of the method, Hoppe-SeyJer's X. Physiol Chem, 1975, 356:1209～1223

42 Snyder RS, Rhodes PH. Electrophoresis experiments in space, in Frontiers in Bioprocessing Sikdar, S. K., Bier, M. and Todd, P., Ed., CRC Press, Boca Raton, FL, 1988. chap. 21

43 Morrison DR, Cell Electrophoresis in Microgravity: Past and Future. In Bauer J ed. Cell Electrophoresis. CRC Press, Boca Raton, FL, 1994

44 Weber G, Bauer J. Counterbalancing hydrodynamic sample distortion effects increases resolution of free-flow zone electrophoresis. Electrophoresis, 1999, 19:1104～1109

45 Hoffstetter-Kuhn S, Wagner H. Scale-up of free-flow electrophoresis: I. Purification of alcohol dehydrogenase from a crude yeast extract by zone electrophoresis. Electrophoresis, 1990. 11, 45. 1～456

46 Barth F, Bruetter MB, Kessler R *et al*. The use of free flow electrophoresis in the purification of recombinant human tissue plaminogen activator expressed in yeast, Electrophoresis, 1986, 7:372～375

47 Loseva OI, Gavryushkin AV, Osipov V *et al*. Application of free-flow electrophoresis for isolation and purification of proteins and peptides. Electrophoresis, 1998, 19:1127～1134

48 Bier FF, Bettag U, Rheingans T *et al*. Determination of the electrophotretic mobility of chromosomes by free flow electrophoresis I. Morphology and stability. Electrophoresis, 1989, 10:690～697

49 Kobayashi H, Ishii N, Nagaoka S. Bioprocessing in microgravity: free flow electrophoresis of *C. elegans* DNA. J Biotechnol, 1996, 47:367～376

50 Volkl A, Mohr H, Weber G *et al*. Isolation of rat hepatic peroxisomes by means of immune free flow electrophoresis. Electrophoresis, 1997, 18(5):774～780

51 Spaans MC, Tobler M, Ammann RW *et al*. Separation and analysis of pig pancreatic zymogen granules with free flow electrophoresis and lectins. Electrophoresis, 1994. 15(5):572～576

52 Hymer WC, Salada L, Krishman K *et al*. Bioprocessing in microgravity: applications of continuous flow electrophoresis to rat anterior pituitary particles. J Biotechnol, 1996, 47:353～365

53 Bauer J, Weber G. Sodium chloride in preparative free-flow cell electrophoresis: recent developments. Electrophoresis, 1996, 17:526～528

54 Kaneko S, Oshio S, Kobayashi T *et al*. Human X and Y-bearing sperm differ in cell surface sialic acidcontent. Biochem Biophys Res Commun, 1984, 124:950～955

55 Engelmann U, Krassnigg F Schatz H *et al*. Separation of human X and Y spermatozoa by free-flow electrophoresis. Gamete Res, 1988, 19:151～159

56 Valet G, Franz G, Lauf PK, Different red cell populations in newborn, genetically low potassium sheep: relation to hematopoietic, immunologic and physiologic differentiation. J Cell Physiol, 1978, 94:215～227

57 Hoffmann P, Ji H, Moritz RL *et al*. Continuous free-flow electrophoresis separation of cytosolic proteins from the human colon carcinoma cell line LIM 1215: a non two-dimensional gel electrophoresis-based proteome analysis strategy. Proteomics, 2001, 1(7):807～818

58 Bier M. Effective principle for scaleup of electrophoresis. In: Sikdar *et al* ed. Frontiers in Bioprocessing, CRC Press, InC. 1988. 236～243

59 Egen NB, Bliss M, Mayersohn M *et al*. Isolation of monoclonal antibodies to phencyclidine from ascites fluid by preparative isoelectric focusing in the Rotofor. Anal Biochem, 1988, 72(2):488～494

60 蔡嘉坤，邵晓霞，夏其昌. 1990, 863—205 预演报告

61 Ostrem JA, Van-Oosbree TR, Marquez R *et al*. Purification of Sl nuclease from *Asperillus* oryzae by recycling isoelectric focusing.. Electrophoresis, 1990, 11:953～957

62 Nagabhushan TL. Appication of recycling isoelectric focusing for purification of recombinant human leukocytei interfer-

ons. Electrophoresis. 1986, 7:552～557
63 顾增旺，夏其昌. 循环式液相等电聚焦方法及其应用. 高技术通讯，1994，4(2):30～32
64 Todd P, Elsaser W. Nonamphoteric isoelectri focusing: stability of borate-glycerol pH gradient in recycling isoelectric focusing. Electrophoresis, 1990, 11:947～952
65 Righetti PG, Barzaghi B, Luzzana M *et al*. A horizontal apparatus for isoelectric protein purification in a segmented immobilized pH gradient. J Biochem Biophys Methods, 1987, 15(3-4):189～198
66 Nilsson P, Wadstrom T, Vesterberh O. Separation of proteins from carrier ampholytes after isoelectric focusing. Biochim Biophys Acta, 1970, 221(1):146～148

第八章　高效液相色谱

对于经典的液相色谱法，流动相在常压下输送，所用的固定相柱效低，分析周期长。20 世纪 60 年代末，引入了气相色谱法的理论，在技术上采用了高压输液泵、高效固定相和高灵敏度的检测器，克服了经典液相色谱法的缺点，具有分析速度快、分离效率高、检出极限低的特点。因此被称为高压或高效液相色谱法(high performance liquid chromatography, HPLC)。

第一节　高效液相色谱的类型及分离原理

高效液相色谱的固定相可以是吸附剂、化学键合固定相(或在惰性载体表面涂上一层液膜)、离子交换树脂或多孔性凝胶；流动相是各种溶剂。高效液相色谱法按组分在两相间的分离机理分类，主要可分为：

高效液相色谱
- 液-固吸附色谱
- 液-液分配色谱
  - 正相分配色谱
  - 反相分配色谱
- 离子交换色谱
- 凝胶色谱(排阻色谱)
- 亲和色谱

1. 液-固吸附色谱

液-固吸附色谱是以液体为流动相，以固体吸附剂如硅胶、活性炭、氧化铝等作为固定相。其基本原理是一些物质在同一吸附剂上的吸附能力的不同。这种吸附作用的本质，很可能是基质和样品间的非专一的偶极与偶极间的作用。在样品进入色谱柱以后，经展开，吸附能力差的物质迁移得快，吸附性强的物质移动得慢，从而一个混合物中几种物质就可按其吸附性被彼此分开。分离过程是吸附与解吸附的平衡过程，也是溶剂分子与样品中溶质分子在吸附剂表面竞争的结果。这种竞争吸附同时表现在不同溶质分子的不同官能团对吸附剂表面吸附-解吸附作用的强弱上。

2. 液-液分配色谱

液-液分配色谱的流动相为液体，固定相是将固定液涂渍或键合在惰性载体表面上。分配色谱是基于样品中各组分在固定相和流动相之间的分配系数不同而进行分离的。可分为正相分配色谱和反相分配色谱两种。固定相的极性大于流动相的极性，称为正相色谱，这种色谱通常用于分离极性较强的化合物，如酚类、胺类、羟基化合物、氨基酸等，分离顺序是极性较弱的组分保留值小，先流出色谱柱；固定相的极性小于流动相的极性，称为反相色谱(reversed phase chromatography)，它是目前高效液相色谱法中应用最多和最有

效的方法，常用于分离非极性化合物，分离顺序是极性较强的组分保留值小，先流出色谱柱。

反相色谱固定相有硅胶键合相填料、多孔碳固定相、苯乙烯型共聚微球填料、氧化铝或氧化锆等作为基质的涂层填料。其中硅胶键合相填料最为常用，大量的是各种烷基硅烷的化学键合硅胶。烷基链长可以是 C_2、C_4、C_8、C_{16}、C_{18}、C_{22}等，但最常用的是 C_{18}，又称 ODS，即十八烷基硅烷键合硅胶。

用反相色谱分离纯化蛋白质时流动相的选择应满足：在 210～280nm 的紫外吸收要小、溶剂易除去、价格低、毒性小、选择性好等条件。常用的有机溶剂有甲醇、乙腈、丙醇等。常用的离子对试剂有三氟乙酸（TFA）、磷酸、甲酸。常用的流动相组成：A 液：H_2O-0.1％TFA；B 液：H_2O-X-0.1％TFA（X 为甲醇、乙腈、丙醇、异丙醇等，按此顺序流动相的洗脱能力依次增强）。

3. 离子交换色谱

离子交换色谱（ion exchange chromatowraphy，IEC）是目前蛋白质化学中应用最广的方法之一。其原理是利用物质的带电部分与具有相反电荷的离子交换剂的相互作用不同，进而达到分离纯化的目的。因为几乎所有的生物大分子都有极性，都带电荷，所以离子交换色谱广泛应用于生物大分子的分离纯化；又由于离子交换色谱处理样品量大，容易操作，它已成为生物大分子分离纯化的重要方法，在生物大分子的分离纯化中约有 75％的工作采用离子交换色谱法。而且由于价廉和纯化后的蛋白质多保持原有活性，是生物大分子分离和纯化中应用广泛的通用型色谱。

离子交换色谱固定相的基质有硅胶基质和有机聚合物两大类。根据离子交换剂所带电荷性质，这类色谱又可以分为阳离子交换色谱和阴离子交换色谱，其中当离子交换剂带阳离子基团时即可吸附阴离子样品，称为阴离子交换色谱；反之，当离子交换剂带阴离子基团时，可吸附阳离子样品，称为阳离子交换色谱。根据可电离基团的解离度，它们又可分为强的和弱的离子交换色谱。常见的离子交换基团有，弱阴：$-NH_2$、$-NHR-NRR'$；弱阳：$-COOH$；强阴：$-NR_3^+$、$-N^+\equiv$；强阳：$-SO_3H$。

离子交换色谱中所用的流动相为盐的水溶液。依据 pH 和盐浓度的变化与使用固定相的种类，可用 pH 梯度或盐浓度梯度洗脱（盐浓度一般由 0.01～1.0mol/L）。在绝大多数情况下，以 IEC 法进行分离和纯化的蛋白质均保持原有的活性，但是某些特殊情况例外，如：当蛋白质的活性部位位于可解离的基团上时，用 IEC 分离可能失活。考虑到蛋白质的活性时往往较多用弱的离子交换色谱，一些较小和较稳定的分子，如氨基酸和小肽等仍可用强的离子交换色谱。当然，在测定结构时，不必考虑活性，只要能达到分离的目的，可任意选用。

在离子交换中，蛋白在交换剂上的吸附要置换出在交换剂上的配衡离子，使流动相中的 H^+ 离子浓度增大或减小，从而会影响蛋白的电荷大小及其在固定相同流动相之间的分配平衡。为维持流动相中 pH 恒定，应用缓冲溶液。一般推荐的适合的缓冲容量应为 10mmol/L，还要注意缓冲液不应干扰蛋白质的交换过程，同时注意温度的影响，以避免因此影响蛋白质的分离和纯化。

4. 尺寸排阻色谱

尺寸排阻色谱(size exclusion chromatography)又称为凝胶过滤(gel filtration)、凝胶色谱(gel chromatography)、分子筛色谱(molecular sieve chromatography)。这类色谱和吸附或交换无关,基本原理是生物大分子进入凝胶填料后,分子大的不能进入凝胶的小的孔穴,而分子小的可以进入小的凝胶孔穴,所以,分子质量大的先通过凝胶填料,而分子质量小的后通过凝胶填料,从而达到分离纯化的目的。

排阻色谱固定相的基质有硅胶[1]和聚合物[2,3]两种基质。硅胶基质的典型代表是TSK-SW 系列,是在硅胶基质上负载亲水性聚合物层形成的。对于聚合物基质的 SEC 填料,经典的有葡聚糖凝胶、琼脂糖凝胶和聚丙烯酰胺凝胶,但它们都是软凝胶,不能耐受高压,难以实现高效分离。现在常用的聚合物刚性凝胶有两种:①高交联度(>40%)苯乙烯-二乙烯基苯共聚物微球,粒度约为 10μm,孔径分布范围很大,为 10~100nm,耐压达40MPa,最高使用温度低于 150℃;②羟基化聚醚多孔微球,粒度约为 10μm,孔径 5~200nm,耐压 20~30MPa,使用温度 10~40℃。

在用排阻色谱分离蛋白质时,一般用等度方式洗脱,通常采用盐离子浓度在 0.05~0.5mol/L 之间的缓冲溶液。有时为了增加蛋白质的溶解度,也会加入有机溶剂、脲或盐酸胍等试剂。

目前所用的凝胶过滤的载体可以根据它们的孔径分为很多种不同的规格。每种规格都有其特定的分级范围,一旦超出分级范围,不论是其上限还是下限,即使分子质量有大有小,但载体的排阻情况基本相同,就无法分离。为此,在使用这种色谱时,要根据所要纯化的样品的分子质量选择所用的凝胶过滤的基质。凝胶过滤除了可对混合物进行分离外,还能提供样品分子质量的信息。要较精确地定出样品的分子质量,可以用一系列分子质量不同的标准样品。先分别用大于所用载体上限分子(常用的是蓝色葡聚糖,分子质量约 200 万)和小于下限的分子(可以用有紫外吸收的氨基酸或其衍生物,也可用有色的盐类,例如 $NiCl_2$),测得它们的洗脱体积,用以标定凝胶过滤柱的 V_o 和 V_i。然后再测定一系列已知分子质量的标准样品在柱上的洗脱体积 V_e,所得的一系列$(V_e-V_i)/(V_o-V_i)$和标准样品的分子质量(m)的对数成线性关系。测得了待测分子质量的样品的 V_e,再由$(V_e-V_i)/(V_o-V_i)$从标准曲线上可以求得未知样品的分子质量。更为简单的方法是以 $\log m$ 对 V_e 作图,所得的直线也能作为测未知样品的标准曲线。

也可用此法达到脱盐和浓缩大分子质量物质的目的,但是所用的量均不大。

5. 亲和色谱

亲和色谱(affinity chromatography)是一种以样品的生物活性为依据的分离分析方法。生物分子的特点是有其专一的活性。例如:酶分子和底物或抑制剂的专一结合;抗原和抗体的结合;激素和其受体的结合;一些维生素和血浆中的运载蛋白质的结合;凝集素和糖类的结合;核酸和有关的结合蛋白等。如果能将上述的例子中的诸多相互作用体系中的一方连接到基质上,使之固定化,就有可能分离纯化专一作用的另一方。

亲和色谱的固定相由基质、空间臂和配体三部分组成。基质有陶瓷、聚苯乙烯、硅胶等。空间臂的存在是因为基质的几何位阻常常限制了分离物与配体的结合,所以,为了使

配体能与分离物的作用位点起作用，在配体和基质间加入一定长度的有机基团即空间臂。配体通常是酶、抗原、抗体、激素、核酸等生物大分子。

亲和色谱的流动相主要是一些缓冲溶液，常用的有磷酸盐(PBS)、硼酸盐、乙酸盐、柠檬酸盐、Tris-HCl 等缓冲体系。

为了能有效地进行分离分析，在将配体和基质连接时，一定要保留配基的生物活性，即使用的连接方法不能过分剧烈。同样道理，在样品被亲和在固定化配基以后，所用的解吸条件也应是尽可能温和，以免样品，或固定化配体，或这两者失活。

亲和色谱的最大特点是其高效和简便。一旦制备固定化配体后，操作使用非常方便，通常情况只要在上样后，洗去杂质，就可将所需的物质解吸下来，而且所得的物质已有较高的纯度。这样的方法可以从含量极低的原料中经一次亲和色谱就能分得所要的物质，得率也较高。以血浆中的维生素 B-12 结合蛋白为例[4]，这种结合蛋白在血浆中的含量仅 5/1 000 万，如果应用常规的分离纯化方法(盐析和离子交换色谱或电泳等)，即使能得到，得率也很低，采用固定化的维生素 B-12，能很方便地纯化得这一结合蛋白。亲和色谱的过程也是高度浓缩的过程。亲和色谱还有两个另外优点，是其他的分离纯化方法所不能比拟的。一是可以从样品中除去大量的杂质，产率较高的得到少量的物质；二是可以在活性物质中除去理化性质几乎完全相同的但已失活了的那些“杂质”，这一点对提高酶或其他分子的比活特别有用。

在亲和色谱中还有三种通用性较强的技术，免疫亲和色谱[5,6]、凝集素亲和色谱[7,8]和染料亲合色谱[9,10]。用蛋白质抗体耦联到载体所形成的免疫亲和柱可用于作为抗原蛋白的分离和纯化。凝集素是一大类能专一地与糖类及糖复合物结合的蛋白质，固定化的凝集素则可用于糖蛋白的分离纯化。

6. 疏水作用色谱

疏水作用色谱(hydrophobic interaction chromatography)是利用在高的盐浓度时，蛋白质会吸附在疏水色谱填料上，当降低洗脱液的盐浓度时，它们会按疏水性的强弱先后被洗脱。因疏水作用色谱有很高的分辨率，所以是生物大分子分离纯化最有效的分离手段之一。因其廉价和纯化后的蛋白具有生物活性而成为一种较新的发展前途广阔的专门用于分离生物大分子的色谱方法，并且是一种通用型的分离和纯化蛋白的工具。

疏水作用色谱的原理与反相色谱相同，区别在于疏水色谱固定相表面的疏水性没有反相色谱强，所用基质分为有机聚合物(交联琼脂糖 Superose 12、TSK-PW、乙烯聚合物等)和大孔硅胶键合相两类。疏水配基有苯基、戊基、丁基、丙基、羟丙基、乙基、甲基以及硅胶表面键合聚乙二醇等。

流动相一般为 pH6～8 的盐水体系，$(NH_4)_2SO_4$、NH_4Ac、NaCl 和磷酸盐等都是常用的盐，其中$(NH_4)_2SO_4$ 应用最为广泛。洗脱方式采用梯度洗脱逐渐减少流动相中盐的浓度，使蛋白质与填料间的疏水作用减弱而使蛋白质得到洗脱。和普通反相色谱相比，疏水作用色谱的疏水性小，且流动相中不采用有机溶剂，蛋白质变性的可能性较小，有利于蛋白质保持固有的活性。

第二节　高效液相色谱的装置

图 8.1 是高效液相色谱仪的结构示意图。主要包括输液系统、进样系统、分离系统和检测系统。工作时，高压泵将贮液池中的流动相输入色谱柱，然后从检测器出口流出。待测组分注入进样器后，被流经其中的流动相带入色谱柱进行分离。分离后的各组分依次进入检测器，检测器的响应信号由记录仪（或装有色谱工作软件的电脑）记录下来即得到液相色谱图。

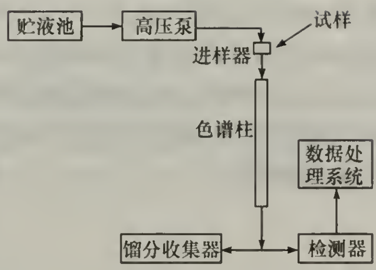


图 8.1　高效液相色谱仪结构示意图

一、输液系统

1. 泵

泵是输送流动相液体的动力，人们希望能得到稳态的连续流动的液体，低流速时的噪声尽量小，所以对泵一般有以下要求。①流量稳定。这一点非常重要，因为流量的稳定性直接与保留时间的准确性相关。对于 4～5mm 的常规色谱柱，最常用的流量范围是0.5～2.0ml/min，误差应不大于 1%。②耐高压。③耐腐蚀。④操作和检修方便，特别是流量调节、高压密封和单向阀的清洗和更换等，要求简便易行。

早期在 HPLC 中大量采用的是往复式的活塞泵，有的公司采用低扩散泵，Isco 和 ABI 公司则采用注射泵（syringe pump），像针筒一样一次推进，这样没有脉冲（图 8.2）。如果一个泵是 20ml，二个泵形成梯度，总体积就有 40ml。

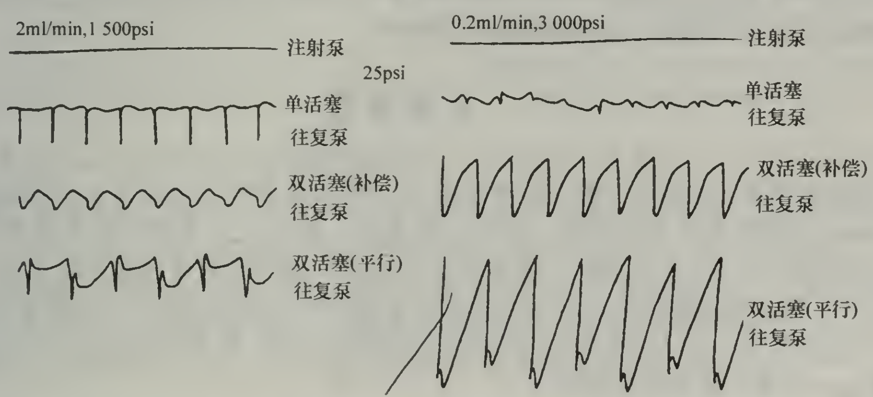


图 8.2　不同泵造成的脉冲

2. 梯度洗脱

当样品混合物的容量因子 k'范围很宽，当等度（isocratic）洗脱时不仅耗费时间长而且后面的峰太扁平不便检测，这时就需要梯度（gradient）洗脱。梯度方式可以是连续的，也

可以是分段式的。梯度洗脱可以使分析时间缩短,使所有峰都处于最佳分离状态,而且峰形比较尖锐。

梯度洗脱时溶剂系统可以是二元、三元,甚至四元的,但最常用的是二元溶剂体系。所用设备也有好几种,大致可分为两类:一类是高压梯度(或称内梯度),即按预先设计的程序分别用两台高压泵把两种溶剂打入一个混合器,混合均匀后再进入色谱柱,此法需两台泵工作,价格较高,但有流量精度高和梯度洗脱曲线重现性好的特点;另一类是低压梯度(或称外梯度),即两种溶剂按一定程序混合,随后由一台高压泵打入色谱柱,此法仅需一台泵,价格便宜。

二、进样系统

HPLC 中的进样方式有注射器进样、停流进样、阀进样、自动进样器进样等。

(1) 注射器进样

即像气相色谱那样,用微量进样器针头穿过橡皮隔膜进样,这是最简便的一种进样方式。而且由于可以把样品直接送到柱头中心,死体积几乎为零,所以往往可获得最好的柱效。

(2) 停流进样

为了避免高压下操作,当使用注射器隔膜进样时可以停泵断流,泄压后注入样品,即停流进样,随后再开泵进行色谱分离。

(3) 阀进样

虽然由于阀接头和连接管死体积的存在,柱效稍低于注射器进样,但因耐高压、重复性好、操作方便而深受色谱工作者的欢迎,其中六通进样阀最为常用。

(4) 自动进样器进样

自动进样器多用于洗脱条件相同且样品量较多时或无人看管的色谱仪。使用微处理机或电脑来控制六通阀的采样、进样和清洗等动作。

三、分离系统

作为分离系统的色谱柱一般为不锈钢管,柱内壁要求精细的抛光加工。分析柱内径为 2～6mm,排阻色谱柱内径粗些。柱长一般为 10～50cm。色谱柱的装填常用两种方法。对于直径大于 20μm 的填料,一般采用干式装柱法,即将填料通过漏斗加入到垂直放置的柱管中,同时敲打或振动柱管,以得到填充紧密而均匀的色谱柱。另一种是匀浆法,此法用于直径小于 20μm 的填料,将填料用溶剂配制成匀浆,然后用加压介质(己烷或甲醇等)在高压下将匀浆压入柱管中。

有人形容,柱是 HPLC 的"心脏",可见在 HPLC 工作中选择柱的重要性。因为用不同的柱,因其分离机理也不同,得到的效果也不同。通常有反相 HPLC、疏水 HPLC、凝胶过滤 HPLC、各种离子交换 HPLC、亲和 HPLC 等。

1. 柱性能的判别

对色谱工作者来讲,柱效是不离口的。而对一般 HPLC 用户来讲,对此并无概念。

在生化界，人们在进行 SDS-PAGE 分析时，常加已知分子质量的标准蛋白质在一侧为对照，已是一种常规程序；而在进行 HPLC 时，很少有人先去检查柱的分辨率和计算柱效，因此在出了问题或感到有问题时非常迷惑，不知如何着手检查原因和排除故障。

新柱到货后，厂方都应有检查柱效的结果报告附上，用户可以据此检查验收，但由于试剂、水质、仪器型号等方面的差异和运输过程的影响，按常规，柱效允许有 10% 的差别。

柱的分辨率和柱效的测定，一般采用几种化合物为标准在特定的条件下进行分离，并据此结果计算柱效[11]。参见图 8.3 和表 8.1 以 Beckman 公司的反相柱为例，采用一些小分子化合物，配伍如下：乙酰苯 33mg，硝基苯 59mg，苯 2g，甲苯 3g，溶于 3L 甲醇中备用。进样 5μl，254nm 检测[11]。

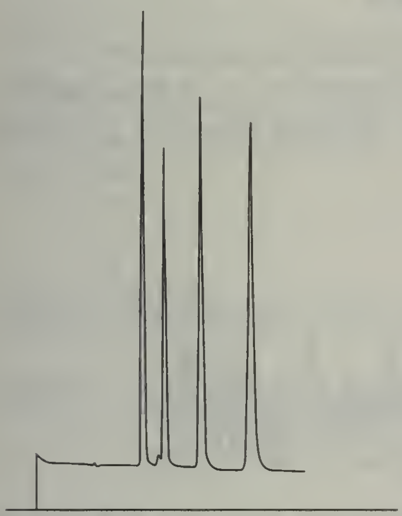


图 8.3 柱效测定和对称性测定
柱：LiChrosorb RP-8，250mm×4.6mm；
洗脱液：ACN/H_2O 75/25
进样：10μl，纸速：2.0mm/s，压力：23.0bar，
检出：254nm，0.25AU 100mV FS

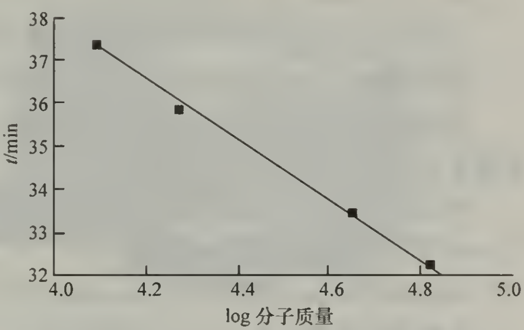


图 8.4 TSK 3000PW 的凝胶过滤测定蛋白质分子质量
样品：细胞色素 c，肌红蛋白，卵清蛋白和 BSA；
流动相：20mmol/L PBS，0.3mol/L NaCl，pH7.0；
流速：0.2ml/min，检测：280nm，40℃

在鉴定 Phenyl 5PW 柱时用丙酮为标准品，水作流动相；在鉴定 CAA-HIC 柱时用 α-胰蛋白酶原为标准品（5mg/ml，进样 5μl），3.0mol/L $(NH_4)_2SO_4$/0.5mol/L NH_4Ac，pH6.0 和 0.5mol/L NH_4OAc，pH6.0，20min 内线性梯度洗脱；在鉴定 DEAE 柱和 CM 柱时，分别用尿酸（0.1mg/ml）和色胺（0.1mg/ml）作标准品；在鉴定 G2000SW 或 G3000SW 时分别用葡聚糖（500 000Da），PEG（7 500Da 和 PEG 3 000Da）或已知分子质量的蛋白质为标准（图 8.4，本实验室结果）。此外还有 Cyano、IP、Si 等柱，都可在厂商的资料中查阅到有关信息。

表 8.1 柱效和对称性等测定

组成/100ml	苯 40.00	萘 3.70	蒽 0.19	1,2-苯并蒽 0.95mg
保留时间/s	223.9	268.7	348.1	445.4
容量因子	0.8	1.1	1.8	2.6
峰对称性	1.2	1.2	1.2	1.2
HETP/μm	28.6	27.9	29.3	31.1
理论塔板数/m	35020.5	35879.1	34147.4	32128.0

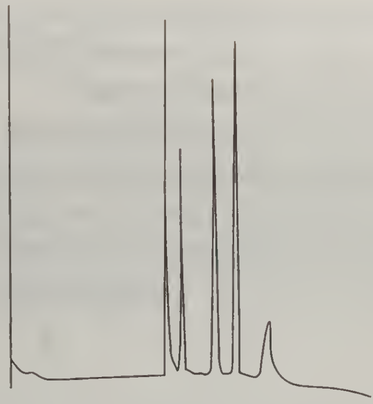


图 8.5 5 种标准蛋白质的分离图谱

ABI RP-300(C8 柱)30mm×2.1mm

以应用生命系统(applied biosystem)的微柱为代表,直接采用一些标准蛋白质来鉴定反相柱、离子交换柱、疏水柱的分离效果,这对生物化学工作者来讲,更加直观。图 8.5 为本实验室结果。

这 5 种蛋白质依次是:细胞色素 c、胰岛素、乳白蛋白、碳酸酐酶、卵白蛋白。有时也用核糖核酸酶、溶菌酶、胰凝乳蛋白酶原等。β-乳球蛋白的胰蛋白酶酶解液(多肽)也用作鉴定柱效。

2. 柱的新进展

(1) 微柱(narrow column, microcolumn)[12]

经典的柱是 250mm×4.6mm,离子交换柱是 75mm×7.5mm。而微柱内径是 1~2mm,长度为 30、100、150、150mm 等。微柱的优点如下所述。

分辨率至少和经典柱一样,长的柱效果更好。

回收率高。因为柱的载体少,非专一性吸附少。

灵敏度高。由于流速仅 100~200μl/min(经典柱 1ml/min),样品的峰尖窄,所以检出灵敏度提高。

溶剂消耗少,节省运转费用。以流速 100μl/min,一个流程 1h 计算,分析一个样品只需 6ml 洗脱液,约仅消耗 3ml 乙腈。

(2) 无孔径载体柱(non-porous resin column)[13]

这种载体颗粒度为 3~5μm,表面无孔径,因此被分析的物质不会进入载体内部,而只在载体表面进行吸附、疏水结合或交换等,故可用于蛋白质和多肽的快速分析,一般流速 1.5ml/min,只需几分钟即可完成[13]。当应用经典的柱分离生物大分子时,通常使用较慢的流速,因为大分子进出柱介质孔径结构的速度较慢,如果为缩短时间而采用较快的流速,则分辨率下降。无孔径载体柱的发展克服了因蛋白质大分子进入载体颗粒内部而引起的局限性,解决了高流速时样品峰太宽及拖尾等问题。

(3) 灌注色谱(perfusion chromatography)[14]

其载体颗粒结构中含有两类孔径(6 000~80 000Å 的贯穿孔和 800~1 500Å 扩散孔),克服了传统填料的传质瓶颈,大大提高了传质效率,使分离线速度从传统的 50~360cm/h 提高到1 000~7 000cm/h。它比常规的 HPLC 柱载体的分离速度快 10 倍,30s~3min 即可完成高分辨率、高载量的分离,是大量制备蛋白质类产品很有潜力的方法(图 8.6)。

(4) 毛细管 HPLC[15]

从 1986 年发展了微柱 HPLC 分离蛋白质和多肽,将 HPLC 流动相流速从经典柱(250mm×4.6mm I.D.)的 1ml/min 降低到 50μl/min(用两个 10ml 的注射泵产生梯度,这类泵的优点是没有脉冲,产生的梯度平稳)。1995 年 Yamada 发展了毛细管 HPLC,用 2.5ml 的注射泵,流速仅 5μl/min;样品上样量 5~25μl;柱为 150mm×0.5mm I.D.;梯度

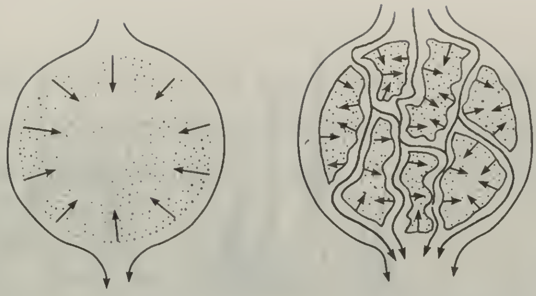


图 8.6　灌注色谱与经典色谱的比较示意图

时变化小于 10nl/min。收集系统改为毛细管直接点涂在 PVDF 膜上，极其方便地定位峰和收集样品(图 8.7)。如此收集在膜上的样品，可用刀片割成 1～2mm，然后放在蛋白质序列仪的反应器上进行序列分析，Hsi 等用于 t-PA 的研究工作中，分析量在 300fmol～10pmol 之间。



图 8.7　毛细管 HPLC 示意图

从经典的柱(内径 4.6mm)到微柱(内径约 2mm)，再从微柱到毛细管柱(内径＜0.5mm)。一般讲，内径小了一半，灵敏度可提高 4 倍(如从经典柱到微柱)。因此，毛细管 HPLC 的灵敏度比经典柱有了很大的提高。此外，由于用小内径的柱，流速也相应降低，又使灵敏度进一步提高(表 8.2)。

表 8.2　柱内径、流速和灵敏度的关系

柱/mm	流速/(μl/min)	峰体积/μl	名义上的浓度
100×4.6	1000	360	1.0
100×2.1	250	75	4.8
100×1.0	50	17	21.2
100×0.32	5	1.7	206.6

实际上，近年来已有十多家厂商生产毛细管柱，一般内径是 0.32mm，长度 150～250mm，颗粒度 3～5μ，品种有正相，反相有 C1(SAS)、C4(Butyl-)、C8(MOS)、C18(ODS)等。图 8.8 是肽图分析图。

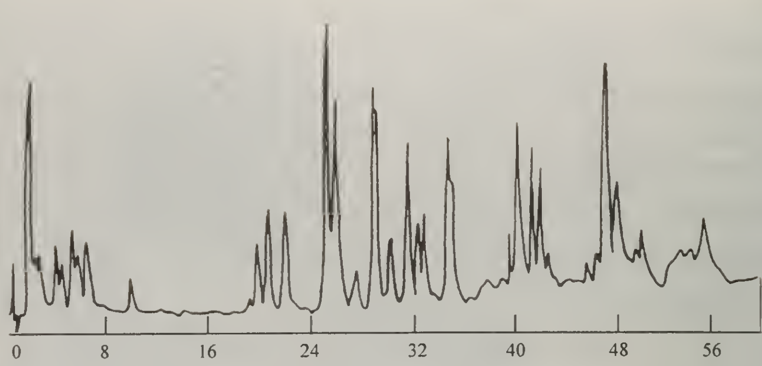


图 8.8 肌红蛋白的胰蛋白酶酶解肽谱

柱:Selectosil C18,5μ 300A,250mm×0.32mm,

流动相:A:2%ACN/98%H_2O/0.1%TFA,B:90% ACN/10%H_2O/0.09%TFA

流速:10μl/min,检出:214nm,进样:5μl,5pmol/μl

四、检测系统

用于 HPLC 的检测器,除应该具有灵敏度高、噪声低、线性范围宽、响应快、死体积小等特点外,还应该对温度和流速的变化不敏感。

高效液相色谱常用的检测器有紫外检测器、示差折光检测器、荧光检测器和电导检测器。

1. 紫外检测器

该检测器适用于对紫外光(或可见光)有吸收性能样品的检测。特点是:使用面广(如蛋白质、核酸、氨基酸、核苷酸、多肽、激素等均可使用);灵敏度高(检测下限为 10^{-10}g/ml);线性范围宽;对温度和流速变化不敏感;可检测梯度溶液洗脱的样品。

蛋白质中有一些芳香族氨基酸,它们在近紫外有吸收光谱,常用的检测波长是 280nm 和 254nm;在远紫外区肽键有吸收,如 210nm,甚至 190nm,在该区域中灵敏度很高,但选择性较差,因为糖类等其他物质在远紫外也有吸收。在一些复合蛋白中,一些蛋白质以外的其他组分也可作为蛋白质检测的依据,例如血红蛋白中的血色素,以及转铁蛋白配位结合的铁离子。

另外,紫外光电二极管阵列检测器也已研制出来,它的基本特点是一次色谱操作中可同时获得吸光度、时间和各组分 UV 光谱图一起的三维谱图,便于得到最佳检测波长、色谱图、光谱图等;并且在一张三维谱上,同时可以得到定性、定量及色谱峰是否是单一组分的信息。

2. 示差折光检测器

凡具有与流动相折光率不同的样品组分,均可使用示差折光检测器检测。目前,糖类化合物的检测大多使用此检测系统。这一系统通用性强、操作简单,但灵敏度低(检测下限为 10^{-7}g/ml),流动相的变化会引起折光率的变化,因此,它既不适用于痕量分析,也不适用于梯度洗脱样品的检测。

3. 荧光检测器

凡具有荧光的物质，在一定条件下，其发射光的荧光强度与物质的浓度成正比。因此，这一检测器只适用于具有荧光的有机化合物（如多环芳烃、氨基酸、胺类、维生素和某些蛋白质等）的测定，其灵敏度很高（检测下限为 $10^{-12} \sim 10^{-14}$g/ml），痕量分析和梯度洗脱样品的检测均可采用。

4. 电导检测器

电导检测器是离子色谱中使用最广泛的检测器。其作用原理是用两个相对电极测量水溶液中离子型溶质的电导，由电导的变化确定淋洗液中溶质的浓度。此检测器死体积小，灵敏度可达 10^{-9}g/ml。

第三节　各类高效液相色谱应用实例

HPLC 用于蛋白质的分离出现在 19 世纪 70 年代末期，二十多年来，HPLC 已成为分离纯化蛋白质非常有效的方法之一。采用的 HPLC 法有反相色谱、离子交换色谱、疏水作用色谱、排阻色谱、亲和色谱等。

用于蛋白质分离的 HPLC 系统应具有梯度洗脱功能，采用紫外或荧光检测，常用的紫外检测波长为 280nm、254nm 或 215nm；荧光检测要求多数蛋白质应在检测之前进行衍生化反应。

一、反相高效液相色谱

Alan[16]等使用反相高效液相色谱对鲁米那肠促胰酶肽释放因子（luminal cholecystokinin-releasing factor，LCRF）进行纯化。首先将样品以 1∶5 的比例用 0.1％的三氟醋酸（TFA）稀释，接着上样至已用 0.1％的 TFA 平衡好的反相 C18 柱（Vydac）上，梯度洗脱。如图 8.9 所示，为检测波长为 220nm 时的色谱峰。

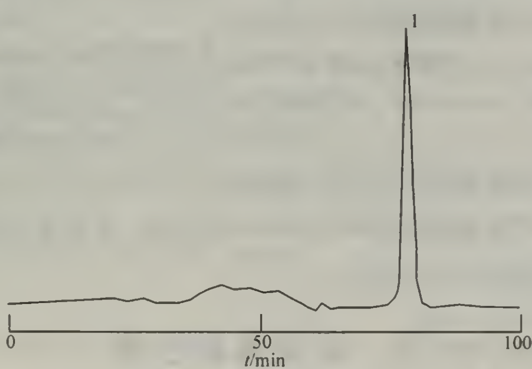


图 8.9　反相高效液相色谱纯化 LCRF 的色谱图
色谱柱：C18 柱（Vydac）；流动相：A：0.1％TFA，
B：0.1％TFA＋乙腈；
梯度：100min 由 100％A→50％B，检测波长 220nm

二、高效离子交换色谱

Laurence 和 Rashmi[17]将大肠杆菌发酵所得到的 MP(movement protein)蛋白的包涵体,使用阴离子交换色谱纯化。将用 8mol/L 尿素溶解的包涵体先用色谱缓冲液(10mmol/L Tris,pH 9.0,8mol/L 尿素,5mmol/L EDTA,150mmol/L NaCl,1mmol/L 2-巯基乙醇)透析,转速为 15 000g 离心 30min,弃沉淀,取上清,上样至 HQ/M 阴离子交换色谱(PerSeptive Biosystems, Framingham, MA)。流动相 A 为上述色谱缓冲液,流动相 B 为 A + 3mol/L NaCl。线性梯度从第 50～250s, NaCl 浓度由 0.15mol/L 增大到 1.15mol/L。

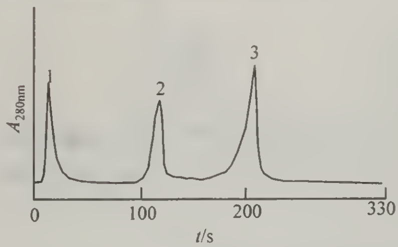


图 8.10 高效离子交换色谱纯化 MP 的色谱图

色谱图如图 8.10 所示。峰 1 为不保留峰,2、3 为当 NaCl 浓度增大时洗脱的峰,经检测,峰 2 和峰 3 为高度纯化的 MP 峰,二者构象上略有差异。

三、高效疏水作用色谱

图 8.11 是高效疏水色谱法分离 7 种标准蛋白的色谱图[18]。色谱柱为 100mm×4mm 不锈钢管,用匀浆法装填西北大学现代分离科学研究所合成的 LHIC-3 疏水填料。使用的是 LC-10A 高效液相色谱仪(Shimadzu),包括主控器(SCL-10AVP)、检测器(SPD-10AVP)、泵(LC-10ATVP)、恒温柱箱(CTO-10ASVP)、Class VP5.03 色谱工作站。

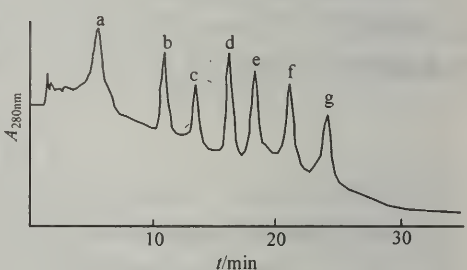


图 8.11 高效疏水色谱柱分离标准蛋白的色谱图

a. Cyt-c; b. Myo; c. RNase-A; d. Lys; e. α-Chy; f. α-Amy; g. Ins

流动相 A 液为 2.5mol/L $(NH_4)_2SO_4$ + 0.05mol/L KH_2PO_4(pH 7.0),B 液为 0.05mol/L KH_2PO_4(pH 7.0),25min 线性梯度(0% B～100% B)洗脱,流速 1ml/min,检测波长 280nm

7 种标准蛋白分别是:细胞色素 c(Cyt-c)、肌红蛋白(Myoglobin)、核糖核酸酶-A(RNase-A)、溶菌酶(Lysozyme)、α-糜蛋白酶(α-Chymotrypsin)、α-淀粉酶(α-Amylose)、胰岛素(Insulin)。

疏水色谱还被用来分离纯化 *E. coli* 表达的 rhIFN-γ[19],rhIFN-γ 的盐酸胍提取液在 40min 内只经一次色谱过程就可以使纯度达到 85%,活性回收率为稀释法的 2～3 倍。

四、高效排阻色谱

排阻色谱除了被广泛应用于蛋白质的分离外,还用于未知蛋白分子质量的测定。图 8.12 为 TSK 3 000PW 的排阻色谱测定蛋白质分子质量。

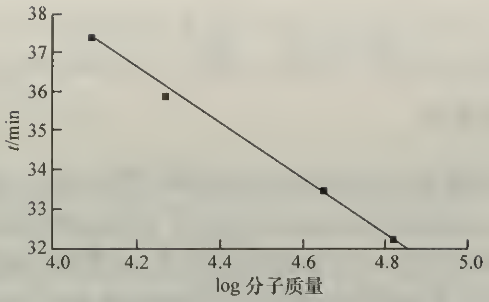


图 8.12　TSK 3 000PW 的排阻色谱测定蛋白质分子质量

样品：细胞色素 c，肌红蛋白，卵清蛋白，BSA

流动相：20mol/L PBS，0.3mol/L NaCl，pH7.0，

流速：0.2ml/min，检测：280nm，40℃

五、高效亲和色谱

粒细胞集落刺激因子（granulocyte colony stimulating factors，G-CSF）是能够在体内外特异地作用于中性粒细胞系，促进其增殖、分化并能维持功能和存活的一种糖蛋白类生长因子[20]。G-CSF 在临床上有非常广泛的用途，其毒副作用很低，目前主要用于防治化疗[21]及放疗后引起的骨髓抑制。此外还用于自身骨髓移植、某些类型的白血病、艾滋病引起的白细胞减少和再生障碍性贫血等，均有显著疗效。免疫亲和色谱可用于 G-CSF 的分离和测定，衍生后荧光检测。如采用免疫亲和色谱法测定化疗诱导的中性白细胞减少症的 G-CSF[22]。

①免疫亲和柱的制备：1g 经酸洗的玻璃小球，用 10% 3-氨基-丙基三乙氧基硅烷回流 16h，将羰基二咪唑接于硅烷化的玻璃球表面。用二氧环己烷清洗，通气干燥。将玻璃小球 1g 置于 50mmol/L pH9.0 含 1mg 抗生蛋白链菌素的 pH7.0 磷酸盐缓冲液（PBS）10ml 中，4℃振荡 18h，PBS 清洗数次，4℃保存。之后加入鼠单抗液，4℃过夜，PBS 清洗小球 5 次，加入到 50mm × 4.6mm 生物兼容柱（PEEK）中。

②流动相：0.1mol/L 盐酸甘氨酸、0.01mol/L PBS 溶液，pH7.0。梯度：柱子先用 0.01mol/L pH7.0PBS 溶液平衡 5min，使 G-CSF 结合到固定相表面，之后 5min pH 作线性变化 7.0～1.5，保持 2min。

③样品衍生化。衍生化试剂为 OPA，取衍生液进样。

色谱图如图 8.13 所示，其中 a 即为 G-CSF 色谱峰，t_R 为 7min。

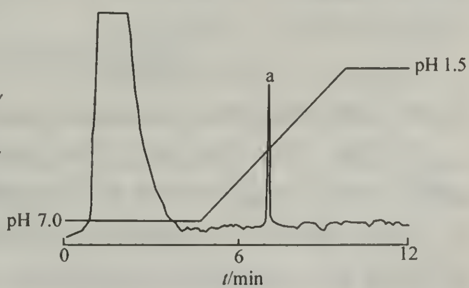


图 8.13　G-CSF 的免疫亲和色谱图

流速：0.5ml/min；温度：4℃；

检测：荧光 340nm/450nm

第四节 讨 论

1. 对蛋白质活性的考虑

所要研究的样品，还要通过活性的检测或其他特性的鉴定，才能从诸多的蛋白质中鉴别分离得到。尽管这一点是显而易见的，但是也常常被从事蛋白质研究不久的生化工作者忽视。为此，不论用何种分法，其中包括色谱技术，分离纯化蛋白质样品前，原则上说，首先建立所要研究的蛋白质的活性检测方法。活性检测方法不仅可以从分离所得到的诸多的蛋白质组分中，鉴别到所要研究的样品，而且还能对所用的分离纯化方法的优劣作出评估。如果在分离纯化后，研究对象的比活（即单位重量的蛋白质所表现的活性）没有提高，则说明研究的对象和其他的杂质蛋白在所用的这种分离过程中没有能分开。如果在分离纯化后，尽管所需的样品比活有所提高，但是总的活性回收率却很低，这提示在整个分离纯化过程中，所研究的样品或是有损失，或是有失活现象。总之，比活提高不明显，或是总活性回收率不高，都说明了所用的分离纯化方法不理想，应该更换其他的分离方法。旨在测定蛋白质一级结构时，可以不考虑所得样品的生物活性，只要能得到某些特定结构或性质的蛋白质，即能用于结构测定。例如可以用专一的抗血清检测，收集具有免疫反应的级分，或是用通过 SDS-PAGE 测定所得蛋白质组分的分子质量，收集所需分子质量的部分。

2. 评价与文献结果的差异

经常有人在重复文献时，得不到相同的结果，甚至差别很大。可能有下述因素要考虑。

①柱效如何。可按第二节的办法去检查。如果柱效很低，再花很多精力在其他条件探讨上也是白搭。

②柱的型号、孔径、颗粒度等是否与文献一样。如反相柱，有 C-18，C-8，C-4，C-2 等亚型。分离氨基酸。肽等则常用 C-18 的柱，孔径 100～300Å；分离蛋白质则常用 C-4，C-8 的柱，孔径 300～1 000Å 为好。

另外，生产厂商也很有关系，由于生产工艺不同[11]，同一厂商的不同批号之间差别不大，而不同厂商的同型号柱（如 C-18）会有较大差别。

③采用不同公司的 HPLC 仪器，由于产生梯度的方式不完全一样，管道的长短和死体积不一样，因此设定的梯度和实际的梯度会不同。即看来两个实验室表面上的条件相同，事实上可能不同。这要求参考文献的试验条件，在具体条件上作一些调整，才能得到满意的结果。盲目相信文献，不如没有文献。

3. 蛋白质和多肽的 HPLC 中的注意事项

(1) *流动相*

对反相 HPLC 而言，最常用的流动相是 0.1％ TFA/ACN 系统，它有低的紫外吸收、低黏度、好的稳定性和分辨率。在 A 液中的 TFA 浓度稍高于 B 液中的 TFA 浓度（0.08％），这样保持基线的平整，不致基线斜度太大。如基线斜率不理想时，可在 A 液中

加 50～150μl 的 20%TFA 来调整基线；如基线下降，则可将 20%TFA 加到 B 液中。

一般讲，肽峰出现的快慢和肽的大小和疏水性有关，如果大肽或肽的疏水性强，则用丙醇或异丙醇代替乙腈作流动相。

(2) 反相柱

一般讲，长的柱和内径细的柱的分辨率和灵敏度好些，常规的柱是 250mm×4.6mm I.D.，我们实验室常用柱是 30mm×2.1mm，100mm×2.1mm I.D. 和 50mm×1mm I.D.，颗粒度 5～7μm，孔径 300Å。对肽的分离常用 C-18 柱，分辨率较高；对纯化重组蛋白质(分子质量在 10 000～20 000Da 之间)，我们常用 C-8 和 C-4 柱，孔径也大一些为好。

(3) 流速

在常规的 250mm×4.6mm I.D. 柱上流速为 1ml/min；在无孔径柱中流速一般是 1.5ml/min；而在 2.1mm 内径的柱上，流速是 200μl/min(50～200μl/min)；1.0mm 内径柱流速 100μl/min。

以微柱而言，流速小时，分辨率高，灵敏度也好些，肽的保留时间(retention time)也相应会增大。

(4) 肽的重演性

不同厂商的 C-18 柱性能有所差异，这是工艺不同引起的。同一种柱在不同的 HPLC 仪器上的图谱也会有些不同，这也是正常的现象，因为不同仪器的管道长短、流动相的混合器、死体积不尽相同等所引起的。此外，每次实验中柱的平衡时间没有严格控制，也会使肽谱的重复性不好。如果柱没有衡温装置，周围环境的温度变化，也会影响肽谱的重复性。

(5) 检出的可靠性

目前很多人用紫外检出的积分面积作为定量的依据，一般也是可以的，但我们要了解这种定量的不充足性。因为通常用 280nm 或 216nm 来分别检出蛋白质或多肽，但不同蛋白质的摩尔吸光系数不同，因此根据积分面积来定量，就可能引进很大的误差。其次，有些杂质可能在 280nm 或 216nm 处没有吸收，或者不是杂质的吸收高峰，这种情况时单用积分面积作为定量的标准，也会产生很大的偏差。

4. 多维色谱在蛋白质组学及蛋白质分离纯化研究中的应用

最传统的蛋白组学研究方法是双向凝胶电泳(2-D SDS-PAGE)和图像分析法。蛋白质在第一向根据等电点的差异分离，在第二向根据分子质量的不同分离。2-D 胶方法已可与 MS 联用，较好地识别胶上各个斑点中的蛋白质。从胶上切下斑点后，蛋白质被消解成肽碎片，然后再用 MS 分析，所得的肽质量用于检索蛋白质数据库以识别蛋白质。

尽管双向电泳是蛋白组学研究的主要方法，但存在耗时多、劳动强度大的缺点。为此，许多科学家正试图简化这一步骤。解决的方法是用多维色谱法，如 Michigan 大学的化学教授 Lubman[23]在分离人乳腺癌全细胞裂解液时，使用一个商品化的蛋白纯化系统先根据等电点的差异来分离蛋白质(chromatofocusing)，该装置有 20 个通道，由膜隔离，加电场后使组分分离。收集馏分，再使用非多孔反相色谱进一步分离纯化，最后上电喷雾 TOF-MS 分析以获得细胞中蛋白质含量。而 Yates[24]等则使用多维液相色谱、串联质谱、数据库搜索等方法对 *Saccharomyces cerevisiae* strain BJ5460 的蛋白质组进行研究，并称

此方法为 MudPIT(multidimensional protein identification technology)法,得到了目前为止最大量的蛋白质组分析数据,分析并检测了 1 484 种蛋白质,其中包括很多低丰度的蛋白质,如转录因子和蛋白质激酶。

在使用多维色谱纯化系统时要仔细考虑柱子的使用顺序,这样可以缩短操作步骤,从而提高总回收率。比如,从离子交换或疏水色谱流出的样品可直接上排阻,之后可直接上反相柱。此方法被用来纯化小鼠唾液腺的上皮生长因子(EGF)[25]、牛脑中的一种 GTPase-activating protein[26]以及从结肠癌细胞系中得到的 A33 抗原[27]。Opiteck 等使用正交的色谱模式:排阻-反相或阳离子交换-反相来分离复杂的蛋白质混合物[28,29]。

5. 联用技术

(1) LC-MS 联用

目前与 LC 联用的 MS 有:快原子轰击 FAB-MS、电喷雾电离 ESI-MS、基质辅助激光解吸电离 MALDI-MS 等,其中 LC-FAB 多用于小的多肽分子的检测,Zang[30]等采用 LC/ESI-MS 联用技术,测定了小鼠血浆和尿液中 ALA 及其代谢产物 CDEPA 和 DEA 的含量。MALDI-MS 是目前灵敏度最高的质谱法之一,Ackermann[31]等用 LC-MALDI 联用测定抗生素混合组分时,可在样品用量极小的情况下,灵敏、准确地得到各组分的分子质量。

(2) LC-NMR 联用

对于生物大分子,较常用的是 LC 与二维 NMR 联用;对更复杂的分子,有更先进的三维和四维异核 NMR,可提供样品的氨基酸组成、蛋白质的局部构象和整体三维构象的信息。Feng[32]等以 HPLC-NMR 联用,测定了 SCH 56592 的 4 种主要降解产物的结构。

(薛卫华　夏其昌)

参考文献

1 Regnier FE, Noel R. Glycerolpropylsilane bonded phases in the steric exclusion chromatography of biological macromolecules. J Chromatogr Sci, 1976, 14:316～320

2 Hjerten S, Yang D. High-performance liquid chromatographic separations on dihydroxyboryl-agarose. J Chromatogr, 1984 316:301～309

3 Hagel L, Lundstrom H, Andersson T *et al*. Properties, in theory and practice, of novel gel filtration media for standard liquid chromatography. J Chromatogr, 1989, 476:329～344

4 Scouten WH. Affinity chromatography bioactive adsorption on inert matrices. New York:John Wiley & Sons, 1981

5 Jack GW. Immunoaffinity chromatography. In:Kenney A, Fowell S, eds. Practical protein chromatography, Totowa: Humana Press, 1992. 125～134

6 Pepper DS. Selection of antibodies for immunoaffintiy chromatography. In:Kenney A, Fowell S, eds. Practical protein chromatography. Totowa:Humana Press, 1992. 135～172

7 王克夷. 亲和层析用于糖蛋白的分离. 见:鲁子贤主编. 蛋白质和酶学研究方法, 北京:科学出版社, 1989. 73～76

8 West IW, Goldring O. Lectin affinity chromatography. In:Kenney A, Fowell S, eds. Practical protein chromatography. Totowa:Humana Press, 1992. 81～90

9 Stellwagen E. Chromatography on immobilized reactive dyes. In:Deutscher MP ed. Guide to protein purification. Meth-

ods in enzymology. Vol. 182. San Diego: Academic Press, Inc. 1990. 343～357
10 Burton SJ. Dye-ligand affinity chromatography. In:Kenney A, Fowell S, eds. Practical protein chromatography, Totowa:Humana Press, 1992. 917
11 夏其昌. 高效液相色谱在蛋白质和酶学中的应用. 见:鲁子贤主编. 蛋白质和酶学研究方法, 北京:科学出版社, 1989. 77～98
12 夏其昌, 屠红旻. 蛋白质和多肽分离分析技术的一些新进展. 生命的化学, 1990, 10:34～37
13 张丽华, 屠红旻, 吴高德等. HPLC无孔径载体柱用于快速分离蛋白质和多肽. 高技术通讯, 1995, 5:45～47
14 Regnier F. Perfusion chromatography. Nature, 1991, 350:634～635
15 Elemer P. Biosystems Repoter, Issue 26, 1995.8
16 Alan W, Spannagel, Gary M *et al*. Purification and characterization of a luminal cholecystokinin-releasing factor from rat intestinal secretion. Proc Natl Acad Sci USA, 1996, 93:4415～4420
17 Laurence MB, Rashmi SN, Theodore WK *et al*. Recombinant tobacco mosaic virus movement protein is an RNA-binding, -helical membrane protein. Proc Natl Acad Sci USA. 2000, 97:7112～7117
18 薛卫华. 西北大学硕士论文, 2001
19 张耀东, 张丽华, 耿信笃. 人干扰素－γ基因工程大肠杆菌的培养及其产物的纯化与复性. 生物工程学报, 1999, 15(2):240～243
20 Zsebo KM, Cohen AM, Murdock DC *et al*. Recombinant human granulocyte colony stimulating factor molecular and biological characterization. Immunobiology, 1986, 172:175～184
21 Bronchud MH, Scarffe JH, Thatcher N *et al*. Phase Ⅰ/Ⅱ study of recombinant human granulocyte colony stimulating factor in patients receiving intensive chemotherapy for small cell lung cancer. Cancer, 1987, 66:809～813
22 Terry M. Immunoaffinity measurement of recombinant granulocyte colony stimulating factor in patients with chemotherapy-induced neutropenia. J Chromatogr B, 1994, 662:307
23 Chong BE, Yan F, Lubman DM *et al*. Chromatofocusing nonporous reversed-phase high-performance liquid chromatography/electrospray ionization time-of-flight mass spectrometry of proteins from human breast cancer whole cell lysates: a novel two-dimensional liquid chromatography/mass spectrometry method. Rapid Commun Mass Spectrom, 2001, 15 (4):291～296
24 Washburn MP, Wolters D, Yates JR 3rd. Large-scale analysis of the yeast proteome by multidimensional protein identification technology. Nat Biotechnol, 2001, 19(3):242～247
25 Nice EC, Fabri L, Hammacher A *et al*. Micropreparative ion exchange high-performance liquid chromatography: applications to microsequence analysis. Biomed Chromatogr, 1993, 7:104～111
26 Nice EC, Fabri L, Hammacher A *et al*. The Purification of a RAP 1 dependent GTP0 ase activating protein from bovine brain cytosol. *J Biol Chem*, 1992, 267:1546～1553
27 Catimel B, Ritter G, Welt S *et al*. Scott AM, Tu GF, Ji H, Moritz RL, Simpson RJ, Burgess AW Purification and characterization of a novel restricted antigen expressed by normal and transformed human colonic epithelium. J Biol Chem, 1996, 271:25 664～25 670
28 Opiteck GJ, Lewis KC, Jorgenson JW. Comprehensive on-line LC/LC/MS of proteins, Anal Chem, 1997, 69:1518～1524
29 Opiteck GJ, Ramirez SM, Jorgenson JW *et al*. Comprehensive two-dimensional high-performance liquid chromatography for the isolation of overexpressed proteins and proteome mapping. Anal Biochem, 1998, 258:349～361
30 Zang LY, Dehaven J, Yocum A *et al*. Determination of alachlor and its metabolites in rat plasma and urine by liquid chromatography-electrospray ionization mass spectrometry. J Chromatogr B Biomed Sci Appl, 2002, 767:93～101
31 Ackermann BL, Gillespie TA, Regg BT *et al*. Application of packed capillary liquid chromatography/mass spectrometry with electrospray ionization to the study of the human biotransformation of the anti-emetic drug dolasetron. J Mass Spectrom, 1996, 31(6):681～689
32 Feng W, Liu H, Chen G *et al*. Structural characterization of the oxidative degradation products of an antifungal agent SCH 56592 by LC-NMR and LC-MS. J Pharm Biomed Anal, 2001, 25:545～557

第九章 质　谱

第一节 引　言

有机质谱仪自诞生以来,一直是有机小分子结构分析的重要工具之一,但它一直不能成为生物学家用以分析生物大分子如蛋白质、多肽、核酸等的有力武器,这是由于有机质谱仪有限的质量范围和电离技术不适合热不稳定、挥发性差的生物大分子的测定。长期以来,质谱研究工作者和生物学家双方都在不断的探索和企盼有朝一日能使有机质谱在生物大分子的分析中有用武之地。在 20 世纪 80 年代末,Karas[1]和 Fenn[2]等先后发明了基质辅助激光电离和电喷雾电离技术,这两个新的电离技术研究的突破性进展,启动了有机质谱在生物领域中的应用,尤其是在蛋白质的结构分析中获得了蓬勃的发展并取得了令人满意的结果,他们不仅用质谱法分析蛋白质的一级结构,即氨基酸的序列分析,而且对蛋白质的高级结构的研究和应用也逐渐广泛起来,如蛋白质非共价复合物的测定[3,4]、H/D 交换法研究蛋白质的高级结构的性质[5]、一定长度的交联剂(cross-linker)[6,7]研究酵母核孔复合体(yeast nuclear pore complex)等,这些研究成果都表明质谱技术在目前已经成为生物大分子结构分析中必不可少的工具之一,有着非常广阔的前景。在开展生物大分子的研究中所应用的质谱仪也是多种多样的,如电喷雾-三级四极质谱仪[8,9]、电喷雾-四极离子阱质谱仪(ESI-quadrupole ion trap MS)[10,11]、基质辅助激光电离-飞行时间质谱仪(MALDI-TOF MS)[1]、基质辅助激光电离-离子阱质谱仪(MALDI-Ion Trap MS)[12,13]、傅里叶回旋共振质谱仪(FTMS)[14,15]、电喷雾-四极-飞行时间质谱仪(Q-TOF MS)[16~18]等。为了使生物学家能够更有效地应用这些仪器开展研究工作,本章就四极杆、四极离子阱、飞行时间、傅里叶回旋共振质谱的基本原理作一些简单介绍。

第二节 四极和离子阱质谱

一、四极质谱仪(quadrupole mass spectrometry)

四极质谱仪结构很简单,它是由两对相互平行的横截面为双曲面或圆形的极杆所组成,见图 9.1。在此极杆上加上两对极性相反的电位 $\pm\phi_0$ 形成四极场。当离子沿 z 轴进入四极场后将受到此电场的作用,在四极场中无空间电荷即理想场情况下,在场中任何一点(x, y, z)的电位可用下式表示:

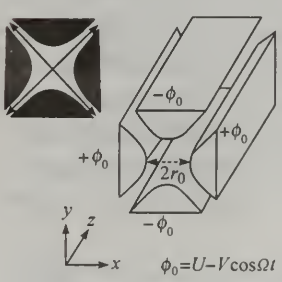


图 9.1　四极质谱仪示意图

$$\phi_{x,y,z} = \frac{\phi_0}{r_0^2}(\lambda x^2 + \sigma y^2 + \gamma z^2) \qquad (9-2-1)$$

ϕ_0 是加在极杆上的电位,

$$\phi_0 = U - V\cos\Omega t \qquad (9-2-2)$$

其中 U 是直流电位，$V\cos\Omega t$ 是射频电位(RF)。V 是射频电位的幅值，Ω 为角频率，$\Omega = 2\pi f$；λ,σ,γ 分别为 x,y,z 方向上的权重常数(weighting constant)。因为场内无空间电荷，电场梯度是均一的，所以满足 Laplace 方程 $\nabla^2\phi=0$，即

$$\nabla^2\phi = \frac{\partial^2\phi}{\partial x^2} + \frac{\partial^2\phi}{\partial y^2} + \frac{\partial^2\phi}{\partial z^2} = 0 \qquad (9-2-3)$$

从(9-2-1)式可得

$$\nabla^2\phi = \frac{\phi_0}{r_0^2}(2\lambda + 2\sigma + 2\gamma) \qquad (9-2-4)$$

要满足(9-2-3)式，则 $\lambda+\sigma+\gamma=0$，或 $\phi_0=0$，这些条件在四极场中必须得到满足。在四极杆质谱中，$\lambda=-\sigma=1,\gamma=0$。在离子阱中，$\lambda=\sigma=1,\gamma=-2$。
在此场内离子在 x,y,z 方向上的运动方程($F=ma$)分别由下式表示

$$F_x = ma_x = m\frac{d^2x}{dt^2} = -e\frac{\partial\phi}{\partial x} \qquad (9-2-5)$$

$$F_y = ma_y = m\frac{d^2y}{dt^2} = -e\frac{\partial\phi}{\partial y} \qquad (9-2-6)$$

$$F_z = ma_z = m\frac{d^2z}{dt^2} = -e\frac{\partial\phi}{\partial z} \qquad (9-2-7)$$

以 x 方向上离子的运动方程为例，对(9-2-1)式微分得

$$\frac{\partial\phi}{\partial x} = \frac{2\lambda x}{r_0^2}(U - V\cos\Omega t) \qquad (9-2-8)$$

把此式代入(9-2-5)式得：

$$F_x = m\frac{d^2x}{dt^2} = -\frac{2\lambda ex}{r_0^2}(U - V\cos\Omega t) \qquad (9-2-9)$$

重新整理方程(9-2-9)可得

$$\frac{d^2x}{dt^2} + \frac{2\lambda ex}{mr_0^2}(U - V\cos\Omega t) = 0 \qquad (9-2-10)$$

令 $\xi=\frac{\Omega t}{2}$，则

$$\frac{d}{dt} = \frac{d\xi}{dt}\frac{d}{d\xi} = \frac{\Omega}{2}\frac{d}{d\xi}, \frac{d^2}{dt^2} = \frac{d\xi}{dt}\frac{d}{d\xi}\left(\frac{d}{dt}\right) = \frac{\Omega^2}{4}\frac{d^2}{d\xi^2} \qquad (9-2-11)$$

将(9-2-11)式代入(9-2-9)式可得：

$$\frac{m\Omega^2}{4}\frac{d^2x}{d\xi^2} = \frac{-2\lambda ex}{r_0^2}(U - V\cos\Omega t) \qquad (9-2-12)$$

两边同乘$\frac{4}{m\Omega^2}$

$$\frac{d^2x}{d\xi^2} = -\frac{8\lambda exU}{mr_0^2\Omega^2} + \frac{8\lambda exV\cos2\xi}{mr_0^2\Omega^2} \qquad (9-2-13)$$

令 $a_x=\frac{8\lambda eU}{mr_0^2\Omega^2}$ $q_x=\frac{4\lambda eV}{mr_0^2\Omega^2}$并代入(9-2-13)式可得：

$$\frac{d^2x}{d\xi^2} + (a_x - 2q_x\cos2\xi)x = 0 \tag{9-2-14}$$

同样,对四极场中任何一个方向的运动方程也可用(9－2－14)式的形式来表示。

$$\frac{d^2u}{d\xi^2} + (a_u - 2q_u\cos2\xi)u = 0 \tag{9-2-15}$$

其中,
$$a_u = \frac{8eU}{mr_0^2\Omega^2} \quad q_u = \frac{4eV}{mr_0^2\Omega^2} \tag{9-2-16}$$

式中 u 代表 x,y,z 坐标,ξ 是无量纲参数,$\xi = \frac{\Omega t}{2}$。(9－2－15)式,即是著名的马绍方程(Mathieu equation),它是描述离子在四极场中运动规律的基本方程。我们所关心的并不是如何解马绍方程,而是离子在场中沿 z 方向注入后能否依次通过四极杆得到分离并检测,即离子在场中的运动轨道是否稳定,它在 x,y 方向上的振幅是否小于极杆之间的间距($2r_0$),亦即由 U,V,m,r_0,Ω 这些参数所决定的 a_u,q_u 值能否使马绍方程有稳定的解。通常马绍方程是否有稳定的解由 a_u,q_u 为坐标的稳定图来表示。图 9.2(a)中的阴影部分为离子在 x 方向上的稳定区,图 9.2(b)中的阴影部分为离子在 y 方向上的稳定区,因为离子只有在 x,y 方向上同时稳定才能使离子顺利通过四极杆。所以,只有 x,y 方向上稳定区的重叠部分即如图 9.2(c)所示的重叠区域 A,B,C,D 才是离子在场中运动的稳定区域。在四极杆质谱中采用接近于原点的 A 区,一般称此区为稳定三角形,其放大图见图 9.2(d)。

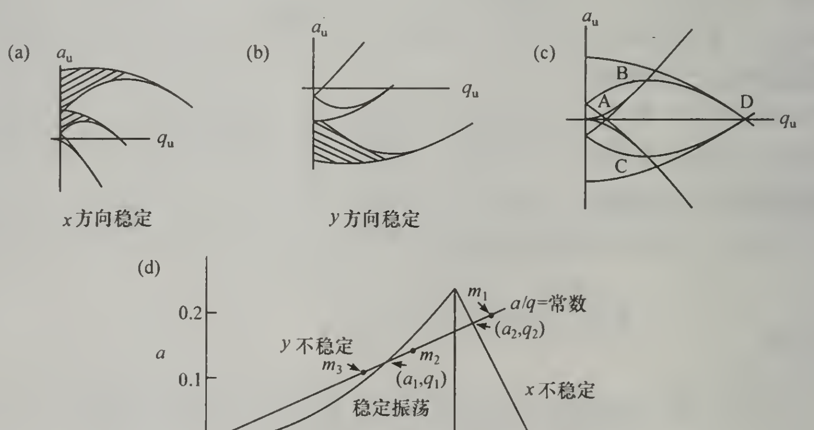


图 9.2 马绍方程稳定图

(a) x 方向稳定;(b) y 方向稳定;(c) x,y 方向稳定重叠区;(d) A 区放大图

图中 a/q 比值为常数并通过原点的直线称为扫描线,此扫描线和稳定三角形边界分别相交于$(a_1, q_1)(a_2, q_2)$,显然,位于此二交点之间线上的离子都是稳定的。从方程(9－2－16)可见 a/q = 常数,即$\frac{U}{V}$ = 常数。若设定一个起始 V 值则 U 值同时被设定,运用(9－2－16)式即可算出不同质量的离子它们此时各自相对应的 a、q 值。从图 9.2(d)中可见只有 m_2,($m_1 > m_2 > m_3$)离子落在稳定区内它可以通过四极场而 m_1、m_3 离子将

与极杆相碰撞而消去。若提高 V 值并保持 U/V = 常数,即起始扫描时,m_1、m_2 将离开稳定区而质量较大的 m_3 将进入稳定区。从以上分析可见,只要保持 U/V = 常数,连续提高射频电位的幅值 V 时,离子将按照其质量的大小,由小到大依次通过四极杆达到检测器,在这种模式操作下,四极场具有质量分离的作用。通常四极质谱仪的质量测量上限范围为 2 000～4 000Da,分辨率为单位质量分辨,属于低分辨仪器。目前在生物大分子,尤其是蛋白质、多肽的结构鉴定中,因要求仪器具有质谱/质谱(MS/MS)功能,一般采用三级四极质谱仪(triple quadrupole mass spectrometer),其结构如图 9.3 所示。

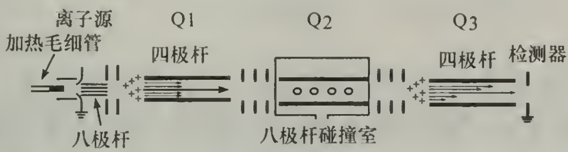


图 9.3　三级四极质谱仪示意图

图中 Q1、Q3 为四极杆,它们分别可单独用作为一个四极质谱仪,Q2 为八极杆,作为碰撞活化室,此八极杆上仅加上幅值较小的射频电位(即 $a_u = 0$),在此条件下,所有质量不同的离子均因它们具有较小的 q 值而能保持在四极场内稳定的振荡,这表明所有由 Q1 选择的离子都能通过 Q2。一旦当选择的离子进入 Q2 后将与 Q2 内充入的惰性气体分子相碰撞,其碰撞能量的一部分转换成离子内能并使其发生裂解,裂解后的碎片由 Q3 作进一步测定,实现了 MS/MS 功能。上述的三级四极质谱 Q1、Q2、Q3 成线形排列。目前商品仪器中亦有成 90°垂直排列的,如图 9.4 所示。这种排列不仅缩小了仪器的尺寸,由于 Q2 成 90°形状,中性碎片不能随着离子流一起转弯到达检测器,所以提高了信噪比。

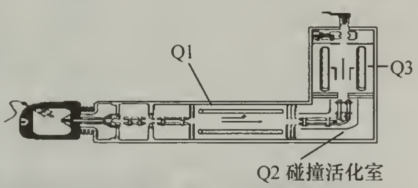


图 9.4　垂直排列三级四极质谱仪

二、四极离子阱质谱(quadrupole ion trap mass spectrometry)

离子阱质谱结构与四极杆质谱相似,由一个环电极(ring)和两个端盖电极(endcap)所组成,它们的横截面均为双曲面,见图 9.5。环电极上的电位为 ϕ_0,端盖电极接地,对于理想的四极场其电极尺寸 $r_0^2 = 2z_0^2$,在此条件下,场中任何一点(x, y, z)上的电位仍可用(9-2-1)式表示 $\phi_{x.y.z} = \dfrac{\phi_0}{r_0^2}(\lambda x^2 + \sigma y^2 + \gamma z^2)$

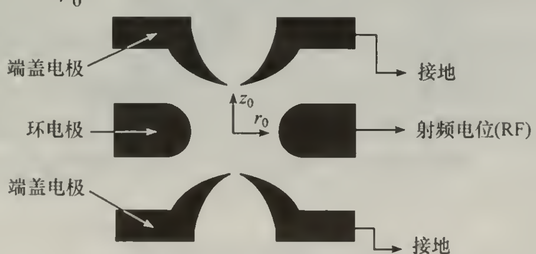


图 9.5　四极离子阱示意图

如前所述，为满足 Laplace 方程的条件，$\lambda=\sigma=1$，$\gamma=-2$，所以四极离子阱电场的表达为

$$\phi_{x.y.z}=\frac{\phi_0}{r_0^2}(x^2+y^2-2z^2) \tag{9-2-17}$$

因为离子阱以 Z 轴为对称轴，所以通常用 r，z 两个方向来表达。

令 $x=r\cos\theta$，$y=r\sin\theta$，$z=z$ 并代入(9-2-17)式，则(9-2-17)式可变换成

$$\phi_{r.z}=\frac{\phi_0}{r_0^2}(r^2\cos^2\theta+r^2\sin^2\theta-2z^2)=\frac{\phi_0}{r_0^2}(r^2-2z^2) \tag{9-2-18}$$

考虑四极离子阱电极上的电位：即环电极($z=0$，$r=r_0$)上的电位 $\phi=\phi_0$ 及端盖电极($r=0$，$z=z_0$)上的电位为 $\phi=0$，所以方程式(9-2-18)按上述边界条件可修正为：

$$\phi_{r.z}=\frac{\phi_0}{2r_0^2}(r^2-2z^2)+\frac{\phi_0}{2} \tag{9-2-19}$$

参照四极质谱中离子运动方程推导过程，亦可得到描述离子阱中离子运动的马绍方程

$$\frac{d^2\mu}{d\xi^2}+(a_u-2q_u\cos2\xi)u=0, \mu(r,z)$$

式中

$$a_r=\frac{4eU}{mr_0^2\Omega^2} \quad q_r=\frac{-4eV}{mr_0^2\Omega^2} \tag{9-2-20}$$

$$a_z=\frac{-8eU}{mr_0^2\Omega^2} \quad q_z=\frac{4eV}{mr_0^2\Omega^2} \tag{9-2-21}$$

与四极质谱一样，由 a_z，q_z 所决定的方程是否有稳定的解，也用稳定图来表示，如图 9.6 所示。如上所述离子阱在正常操作条件下，端盖电极接地，环电极上只有射频电位即 $U=0$，即 $a_z=0$，可见离子的"运动"均在 q_z 轴上进行。设想离子阱在起始时环电极加上射频电位幅值为 V_0，从方程(9-2-21)可计算出不同质量的离子的 q_z 值，凡是其 q_z 值在 0～0.908 之间离子均被稳定地陷在离子阱中，由于 q_z 与 m 成反比，所以质量大的离子 q_z 小，相反亦然。当射频电位的幅值从 V_0 逐渐升高时，不同质量的离子的 q_z 值也随

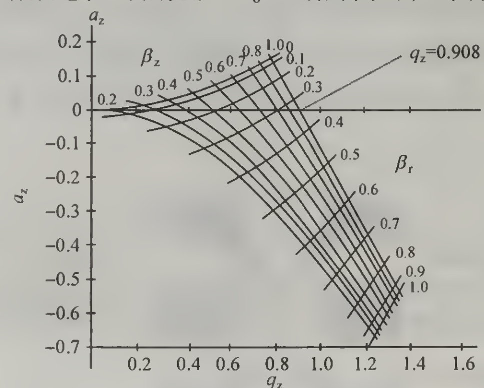


图 9.6　四极离子阱稳定图

之逐渐增大,也就是说它们按其质量从小到大的次序沿着 q_z 轴向右移动,当离子的 q_z 值大于 0.908 时,离子在 z 向的振幅变大,运动变成不稳定并通过端盖电极上的小孔打到倍增器上被检测,以此得到一张完整的质谱图。这种操作模式称为选择质量不稳定模式(mass-selective instability),是离子阱通常使用的操作模式。

离子阱技术从早期离子阱检测器 ITD(ion trap detector)发展到离子阱质谱,(ion trap mass spectrometry)它的主要技术性能如分辨率、质量测量范围和质谱-质谱功能都得到了发展和提高,这也是采取了以下一系列技术措施的结果。

1. 缓冲气(buffer gas)

离子阱充入一定量的惰性气体如氦气(10^{-3}torr)后,离子和氦原子将发生碰撞,使离子的动能得到一定程度的"冷却",从而使离子在 r, z 方向上的位移趋向于离子阱的中心,如图 9.7 所示。当这些离子达到轨道不稳定的临界状态时,它们几乎在离子阱内同样的位置开始向外运动,由于离子不仅成群结队地向外运动而且它们都紧紧地沿着 z 轴方向上聚焦,所以提高了仪器的分辨率和灵敏度。



图 9.7 离子阱中离子在有无缓冲气时的运动状态图

2. 自动增益控制(auto gain control, AGC)

当样品在离子阱中的浓度发生变化时,不论是否有缓冲气的存在,谱形仍会发生变化,致使定量分析无法进行,谱库检索无法实现。这是由于高浓度的样品分子如果在离子阱中陷入较长时间将会发生分子离子反应,从而改变了离子的特征,如分子离子变成了准分子离子,造成定量分析困难。另外,离子阱中离子密度的增加也会引起空间电荷效应(space-charge),由于它改变了"纯"四极场的性质,其结果将造成稳定图边界位置的改变,从而使离子质量发生偏离。为了克服上述缺点,采用了自动增益控制技术,即在扫描过程中设有两个电离时间段,第一电离阶段的时间为固定的,如 0.2ms,电离后赶出全部从背景气中形成的离子(如 m/z 为 44 以下),并测定保留在阱内的离子的总离子流,以此作为依据计算第二阶段电离所需的最佳时间,这样可以避免因阱内离子太多而引起的空间电荷效应。当然在每一次扫描过程中都同时伴有这两个时间段的电离过程,其电离时间与谱的强度亦同时被记录,以便实现数据的归一化处理。由于自动增益控制技术的应用,离子阱质谱也可实现定量分析。

3. 轴向调制(axial modulation)

当离子阱中的射频电位幅值扫描时,低质量离子首先被"扫出"离子阱而较高质量离

子仍旧陷在阱中，由此引起的空间电荷效应将会造成即将被扫出的离子峰的加宽。为了改变这种状况，在端盖电极上加上一个附加振荡电场(oscillation field)，它的峰值为 $6V_{(p-p)}$，其频率约为环电极上射频电位频率的一半。当离子随着射频电位幅值的增加而达到稳定三角形边界($q_z=0.908$)即将被扫出时，在附加振荡电位的作用下，这些离子将产生共振，并使它们紧密地聚集在一起，同时排出离子阱，所以分辨率随之得到提高，如图 9.8 所示。

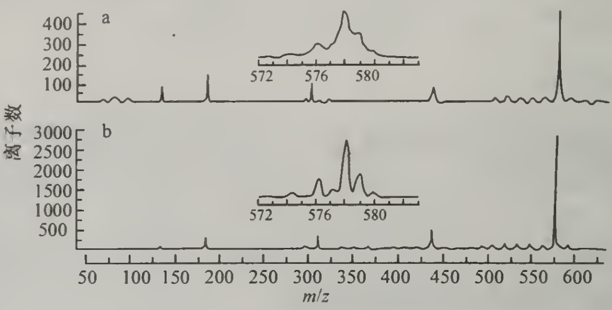


图 9.8　轴向调制效果图

a. 无轴向调制；b. 有轴向调制

4. 质量范围的扩展

离子阱质谱最初设计的质量范围仅为 650Da，显然不适合于生物大分子的分析。从方程 $m/e=\dfrac{4V}{r_0^2\Omega^2q_z}$ 中可见，减少电极尺寸 r_0 或降低射频电位的频率 Ω，均可扩展质量范围，然而最有效的办法是采用共振排斥法(resonance ejection)如图 9.9 所示。在 q_z 轴上设定一个点作为离子共振排斥点，它相当于四极离子阱正常工作时在稳定图的 q_z 轴上开了一个小孔，让离子在此孔的位置上发生共振排斥。其具体操作过程如下：当射频电位幅值扫描时，即在 q_z 扫描时在端盖电极上加上一个频率较低的附加交流电位，它的频率由质量范围扩大的倍数决定。阱内离子随着射频电位的幅值的增大其 q_z 也增大并达到设定的共振排斥点所相应的 q_z 值时，离子依次产生共振并排出离子阱到达检测器。由于 m/z 值与 q_z 成反比，所以共振排斥点的 $q_{z\text{-new}}$ 可用下式表示：$q_{z\text{-new}}=q_{z\text{-eject-old}}(m/z_{\text{lim-old}}/m/z_{\text{lim-new}})$，当欲扩大的质量范围即 $\left(\dfrac{m/z_{\text{lim-old}}}{m/z_{\text{lim-new}}}\right)$ 确定后 $q_{z\text{-new}}$ 即可由上式计算得到，通过 q_z 与 β_z 的关系式(见后)，β_z 即离子阱端盖电极上的附加交流电位的频率也可决定，如需扩大质量范围为 10 倍，则 $q_{z\text{-new}}=0.908/10=0.0908$，相应的 $\beta_z=0.064$，此时在端盖电极上应附加的交流电位的频率为 35.2kHz。

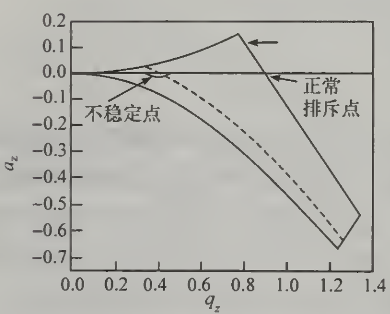


图 9.9　共振排斥法示意图

5. 狭扫描(narrow scan or zoom scan)

通常离子阱的扫描速度为 5 555Da/s,当质量范围为 650Da 时,此扫描速度相当于 180μs/Da。在这种扫描速度下,以质量选择不稳定模式工作时,对单电荷离子其分辨率(半峰宽定义)理论上可达 3×m。离子阱的分辨率除了由它本身的性能决定外,还取决于控制射频电位变化的电子线路中数字模拟转换器(DAC)的速度,在仪器设计中一般将此 DAC 所需的时间间隔等同于离子流信号采集时模拟数字转换器(ADC)的时间间隔,以 12 位(bit)的 DAC 和 ADC 为例,此时间间隔为 28μs,由此可得每一个质量中的采样次数为 180μs/28μs,即约 6.3 次;若采用上述的质量扩展技术,使质量范围扩展至 4 100Da,则每一个质量中的采样点将小于 1。从信息学理论可知,一个峰中采样数至少 6 次才能较完整地表达一个峰的原来面貌。如图 9.10 所示,如果采样次数太少,峰形将发生变化,使原来两个能相互分离的峰合并成一个峰。

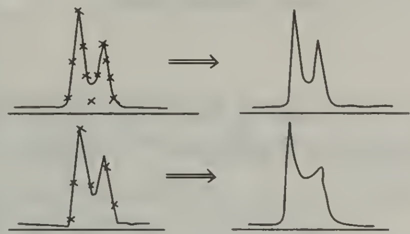


图 9.10　采样次数对分辨率的影响

为了克服由于电子线路上的技术问题所造成对仪器实际分辨率的影响,一种称为补偿 DAC 技术已被广泛应用于离子阱质谱,其工作原理见图 9.11。射频电位扫描是由一个 0～10V 的数字模拟转换信号来控制的,这个电压在其扫描过程中任何时刻均可由 TTL 控制脉冲经电压分配器将它衰减,如图中被衰减为 0～1V,但其扫描速度保持不变。这个衰减的电压经运算放大器与补偿电压(6V)相加后成为射频电位的控制信号,如图中为 6V～7V,这个控制信号经射频电位发生器最终得到射频电位的扫描电压。由此可见,射频电位经此补偿 DAC 修正后扫描速度不变但幅值扫描范围变窄,扫描范围仅从4 850V 到 5 870V,这就意味着在同样的扫描时间内,扫描质量范围变窄,这样就能保证在一个峰中有足够的采样次数,使仪器的分辨率得以真正体现出来。商品仪器在采用此技术后其分辨率将大于 10 000(半峰宽定义)。在蛋白质、多肽样品的分析中,仪器的高分辨能力是至关重要的,因为只有高分辨才能使准分子离子与其同位素峰相分离,图 9.12 为血管紧

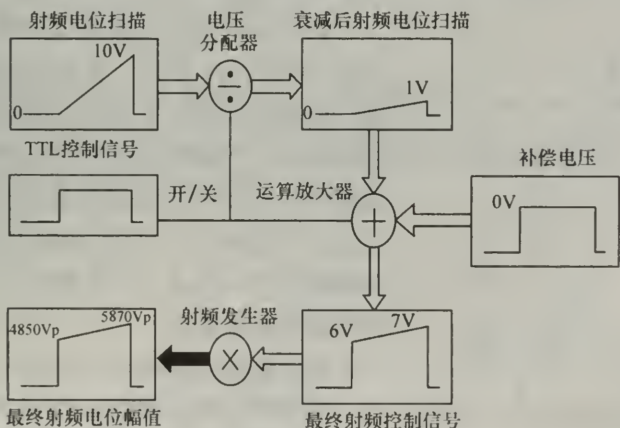


图 9.11　DAC 补偿技术原理图

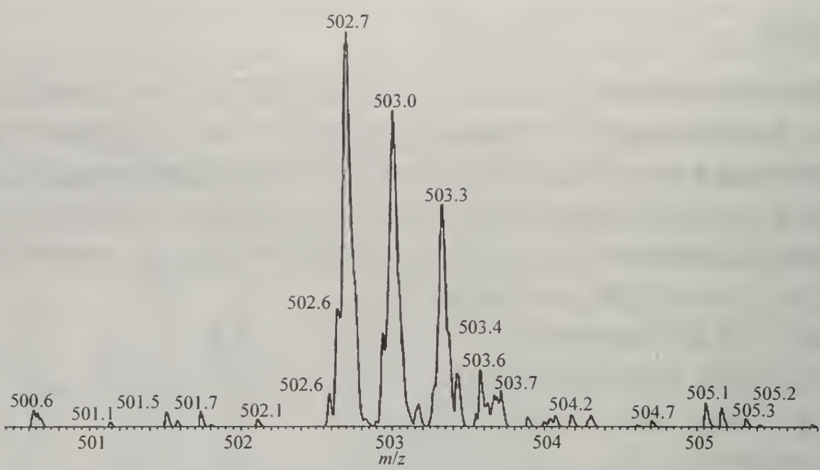


图 9.12　血管紧张素肽段的狭扫描图

张素中某肽段的狭扫描图，图中可见峰间的质量差为 0.3 个质量单位，这表明 m/z 502.7 是一个带有三个电荷的离子。离子所带电荷数的判断，可通过观察谱中一个质量单位内所含有的同位素的峰数来推测，此峰数即为此离子所带的电荷数，从离子的电荷数与质量即可计算出肽段或蛋白质的分子质量。

6. 质谱/质谱(MS^n, $n \geqslant 10$)

在生物大分子分析中，特别是蛋白质、多肽分析中，质谱/质谱功能是必不可少的技术手段。因为肽段的 MS /MS 所提供的各级子离子信息包含了极为丰富的 N 端和 C 端离子结构信息(即 b, y 系列离子)，它不仅为实现计算机数据库检索创造了良好的条件，也为多肽的 *de novo*[19]序列分析提供了所必需的信息。

在离子阱中实现 MS/MS 的关键在于如何提取先驱离子(precusor ion)，即把质量大于或小于此离子的离子赶出离子阱。实现提取单一离子的方法很多，下面仅作简单介绍。

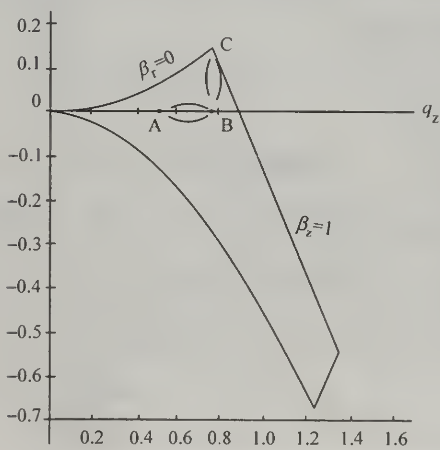


图 9.13　顶点提取法原理

(1) 顶点提取法(apex isolation)

此方法是采取射频电位和直流电位相结合的方法，如图 9.13 所示。首先使提取离子从 A 点移到三角形的顶点在 q_z 轴上的投影点 B，然后在环电极上加上一个负的适当大小的直流电压($-U$)，使该离子的 $q_z>0$，并恰好到达在三角形顶点之下的 C 点。在此条件下，质量比提取离子大或小的离子分别越过 $\beta_r=0$, $\beta_z=1$ 边界线，由于它们在 y 和 z 向不稳定而被消去。

(2) 共振排斥(resonance ejection)

共振排斥法原理如前所述这里不再重复。运用共振排斥法实施离子提取时端盖电极上的附加电场的频率不是固定的，而是以提取离子质量的不

同而改变。通常此频率设定在相当于提取离子质量数加一的水平上。其工作过程如图 9.14 所示，当离子形成后(A)，射频电压立即起始扫描(B)，同时端盖电极加上附加电场，此时小于提取离子质量的离子随着射频电位幅值的增加在 $\beta_z=1$ 边界上，由于 z 向不稳定而被排出离子阱，大于提取离子质量的离子由于与附加电场的频率发生共振而同样被排出离子阱，只有提取离子能留在离子阱中。

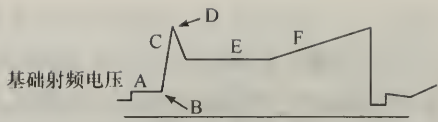


图 9.14　共振排斥法 MS/MS 操作示意图
A. 电离和储存质量范围确定；B. 附加电场开(频率相应于母离子质量 + 1)；C. 提取母离子；D. 附加电场关；E. 碰撞诱导解离；F. 射频电位幅值扫描得到子离子谱

(3) 过滤噪声场法(filtered noise field)

运用此方法时先在环电压上加上射频电位，使一定质量范围内的离子均陷于阱中，其中包括欲提取的离子，与此同时在端盖电极上加上附加振荡电位，其频率覆盖上述质量范围内所有离子的在 z 向的振荡频率而独缺一个相应于提取离子的频率。因此，大于或小于提取离子质量的所有离子均因共振而被排出离子阱，惟有提取离子被留在阱内。

在离子提取技术的发展和应用上，Cooks[20,21]，Kelley[22,23]，Mclucky[24] 等作出了很大的贡献，他们分别采用逆向傅里叶转换的尖脉冲，陷波滤波技术，以及反向-正向扫描等方法，达到了有效提取单一离子并实现了质谱-质谱测试功能，有关这些技术的详细内容在此不一一赘述。

用以上这些方法得到的离子，在阱内经共振激发与背景气体发生相互碰撞并发生单分子裂解反应，然后以通常射频电位幅值的扫描得到子离子图谱。重复上述的过程，即可在离子阱质谱中实现多级 MS/MS 功能。目前商品仪器中最多可达 10～11 级。

在以上讨论离子阱的性能改进过程中，共振排斥、轴向调制技术被广泛应用，这些技术的基本原理均是在端盖电极上加上一个附加的振荡电位，使其与不同离子发生共振，所以，离子的基本频率是离子阱中一个重要参数，它是由下式决定的：

$$\omega_{u,n} = \left(n + \frac{1}{2}\beta_u\right)\Omega \qquad 0 \leqslant n \leqslant \infty$$

$$\omega_{u,n} = -\left(n + \frac{1}{2}\beta_u\right)\Omega \qquad -\infty \leqslant n \leqslant 0 \tag{9-2-22}$$

$$u(r.z)$$

当 $n=0$ 时，称此为离子的基本频率，$\omega_{u,0}=\frac{1}{2}\beta_u\Omega$。

由此可见，凡是 β 相同的离子其基本频率是相同的，这就是稳定图中的等 β 线。β 值是由 a，q 值来决定的，β 值可用下式计算：

$$\beta_u^2 = a_u + \cfrac{q_u^2}{(\beta_u+2)^2 - a_u - \cfrac{q_u^2}{(\beta_u+4)^2 - a_u - \cfrac{q_u^2}{(\beta_u+6)^2 - a_u - \cdots}}} + \cfrac{q_u^2}{(\beta_u-2)^2 - a_u - \cfrac{q_u^2}{(\beta_u-4)^2 - a_u - \cfrac{q_u^2}{(\beta_u-6)^2 - a_u}}} \tag{9-2-23}$$

从(9－2－23)式可见，为计算 β 值，应采用数值逼近法，即逐一代入 β 值，然后比较 β_{n+1} 和 β_n 的差值，直至此差值小至忽略不计为止，在得到 β 值后，显然 ω 值也随之而定，但在离子阱质谱中为提高分辨率、灵敏度，实现质量扩展，提取单一离子所采取的一系列"共振"技术中，如共振排斥、轴向调制等其在端盖电极上附加振荡电位的时间以及幅值是各不相同的，如轴向调制中其频率约为环电极射频电位频率的一半，而幅值大于 $6V_{(p\text{-}p)}$；而在共振激发中(resonance excitation)，其频率为保留在阱中的提取离子所相应的频率，而其幅值不能过大，以能达到提高碰撞效率即可，避免提取离子排出离子阱。通常其幅值小于 $1.5V_{(o\text{-}p)}$。

第三节　飞行时间质谱

Wiley 和 Mclaren 于 1955[25]年发表了他们设计的飞行时间质谱(time-of-flight mass spectrometry，TOFMS)。早期以飞行时间质谱配以激光脉冲电离技术应用于同位素比、元素分析等，直到 1974 年由 Macfarlane[26]实现了等离子体解析电离(plasma desorption)与飞行时间质谱相匹配后才得以在生物大分子的分析中得到应用，尤其是 Karas 和 Hillenkamp 于 1988 年引入了基质辅助激光电离(MALDI)成功地测试了分子质量超过 300×10^3Da 的蛋白质后飞行时间质谱才获得了新生，随着飞行时间质谱性能(质量分辨率、灵敏度等)的不断改进，基质辅助激光电离-飞行时间质谱已成为生物大分子尤其是蛋白质、多肽分析中有效的工具之一。下面对 TOFMS 作一介绍。

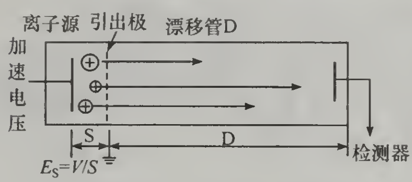


图 9.15　飞行时间质谱示意图

TOFMS 由离子源(S)引出极、漂移区(D)和检测器所组成。如图 9.15 所示。当离子在离子源内引成后在离子源内电场 E 的作用下进入无场漂移区。在理想状态下，所有进入漂移区的离子具有相同的动能(KE)。

$$KE = qE_s \tag{9-3-1}$$

式中 $q=Ze$ 为离子的总电荷，E_s 为加速电压。

$$qEs = \frac{1}{2}mv^2 \tag{9-3-2}$$

式中 m 为离子的质量，v 为离子飞行的速度。

离子在漂移区内所需的时间为

$$t = D/v \tag{9-3-3}$$

以(9－3－2)式中的 v 代入(9－3－3)式可得

$$t = \left(\frac{m}{2ZeEs}\right)^{1/2}D \tag{9-3-4}$$

整理(9－3－4)式为

$$\frac{m}{Z} = 2eEs\left(\frac{t}{D}\right)^2 \tag{9-3-5}$$

或者 $m=At^2$，式中 A 为常数。所以，测定离子在漂移区内的飞行时间即可计算出它的质荷比，亦即离子的飞行时间谱可转换成质谱图。

早期 TOFMS 都是线性飞行时间质谱(linear TOFMS),分辨率较低,这与离子在离子源内形成及加速过程中的离子形成的时间(t)、离子的初始动能(U_0)、离子在离子源中的位置(S_0)及运动方向有关。如图 9.16 所示,图 9.16(a)是两个质量相同、初始动能相同的离子,但是由于它们在源内形成时间的不同($\Delta t \neq 0$),所以当它们以相同速度飞向检测器时,一直保持着它们的时间差。对方程(9-3-5)微分可得

$$\frac{dm}{dt} = 2At, \frac{dm}{m} = \frac{2dt}{t}, \text{所以} \frac{\Delta m}{m} = 2\frac{\Delta t}{t} \tag{9-3-6}$$

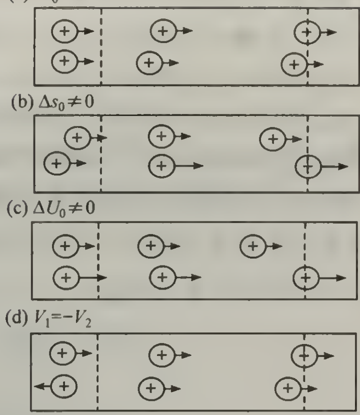


图 9.16 离子源内离子不同状态示意图

显然只要峰宽保持不变(Δt 为常数),延长飞行时间,降低加速电压,可提高分辨率。图 9.16(b)是两个质量相同、初始动能相同,但在源内的不同位置形成的离子($\Delta s \neq 0$)。由于左边的离子比右边的离子受到更强的电场作用,飞行速度较快,相对较早地到达检测器,其结果造成峰宽的加宽。图 9.16(c)是两个质量相同、在同一位置形成,但具有不同初始动能的离子($\Delta U_0 \neq 0$),动能高的离子飞得较快。如果飞行管的长度加长,则这两个离子的飞行时间差也加大,显然不利于分辨率的改善。提高加速电压可减少这种影响。图 9.16(d)是两个质量相同、初始动能相同,但起始飞行方向相反的离子的运动的情况($V_1 = -V_2$)。显然,初始运动方向与飞行方向相反的离子需要更长时间到达检测器。提高加速电压可缩短由此产生的飞行时间差。

综合以上两个质量相同的离子在离子源形成过程中的状态及加速过程中的各种因素的影响,离子在加速电压 Es 的作用下飞越长为 D 的漂移管所需的时间可用下式表示:

$$t = (2m)^{1/2}\frac{[(U_0 + eEs_0)^{1/2} \mu U_0{}^{1/2}]}{eE} + \frac{(2m)^{1/2}D}{2(U_0 + eEs_0)^{1/2}} + t_0 \tag{9-3-7}$$

式中 t_0 为离子形成时间,s_0 为离子在离子源中的位置,U_0 为离子的初始动能。

上式中第一项为离子在电离/加速区内所需的时间。其中$(U_0 + eEs_0)^{1/2}$为离子飞出离子源的速度,$\pm U_0$ 反映了离子初始速度方向的影响。第二项为离子在漂移区内所需的时间,它反映了离子初始动能和离子位置的影响。第三项为离子形成的时间。如果样品分子是从一个表面上经解析电离进入气相的情况,如在 MALDI 条件下,则 eEs_0 为常数,$eEs_0 = eV$(V 为加速电压),在漂移管的尺寸比离子源大得多的条件下($D \gg s$),(9-3-7)式可简化成

$$t = \frac{(2m)^{1/2}D}{2(U_0 + eV)^{1/2}} \tag{9-3-8}$$

我们比较两个动能分别为 0 和 U_0 的离子,它们到达检测器的时间差

$$\Delta t = D\left(\frac{m}{2}\right)^{1/2} \times \left[\frac{1}{eV^{1/2}} - \frac{1}{(eV + U_0)^{1/2}}\right] \tag{9-3-9}$$

从(9-3-9)式可见,离子初始动能的分布(kinetic energy distribution)是影响飞行时

间质谱分辨率的主要因素。把(9－3－6)式代入(9－3－9)式,则分辨率可用下式表示:

$$\frac{\Delta m}{m} \cong 2\left[\left(1+\frac{U_0}{eV}\right)^{1/2}-1\right] \qquad 或者更简化为\frac{\Delta m}{m}\cong\frac{U_0}{eV} \qquad (9-3-10)$$

从(9－3－10)式可见如果提高加速电压或减少初始动能分布将有利于分辨率的提高。为此,Mamyrin[27]和他的同事们于 1973 年在飞行时间质谱中引进了反射器(reflectron),以补偿离子初始能量分布的影响,提高仪器分辨率,其结构见图 9.17。反射器处于

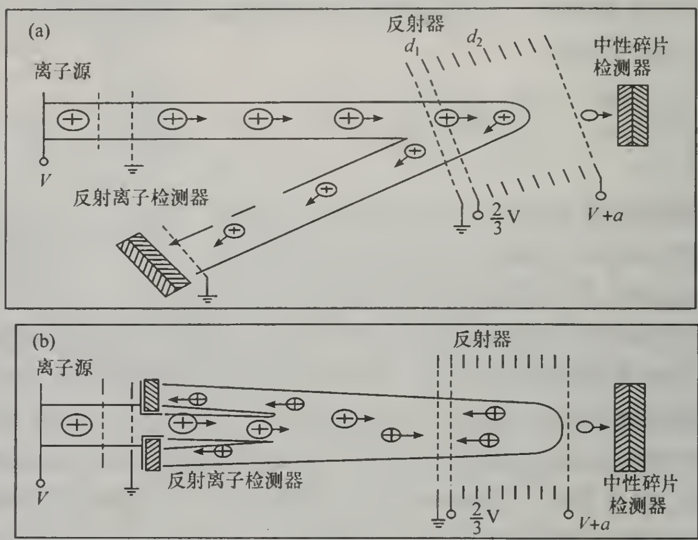


图 9.17　反射器结构示意图

漂移区的后面,里面装有一系列电极,电极上的电压逐极增加,并在最后电极上达到比加速电压略高水平上即($V+a$)。离子穿进反射器后,当它的能量达到 0 后,随即反相加速出反射器,其能量与进入时相同,只是速度方向相反。假定反射器的电场为 E,离子电荷为 q,离子的动能为 K,则它能够穿越的深度为 x,$x=\frac{K}{qE}$。因离子穿进反射器的速度是从 v_{ix}至 0,所以其平均速度为 $v_{ix}/2$。在反射器中飞行长度为 x 所需时间为 $t_0=\frac{x}{v_{ix}/2}$。离子在反射器中穿入及返回的距离为 $2x$。所需的总时间为 $t_r=2t_0=\frac{2x}{v_{ix}/2}=\frac{4x}{v_{ix}}$。对于两个质量相同的离子,若它们具有的动能不同,分别为 K 和 K',且$\frac{K'}{K}=\alpha^2$。因 $K=\frac{mv_{ix}^2}{2}$,$K'=\frac{mv'_{ix}}{2}$,所以它们的飞行速度分别为 $V_{ix}=(2K/m)^{1/2}$;$V'_{ix}=(2K'/m)=(2K\alpha^2/m)^{1/2}=\alpha(2K/m)^{1/2}=\alpha V_{ix}$。以这样的速度,两个离子在漂移区内飞行的时间为 $t=D/V_{ix}$,$t'=D/V'_{ix}=D/\alpha V_{ix}$,$t'=t/\alpha$,而在反射器中飞行 $2x$ 与 $2x'$所需的时间 $t_r=4x/V_{ix}$,$t'_r=4x'/V'_{ix}$。由于这两个离子的能量不同,所以在反射器内穿越的深度也不同,其值分别为 $x=K/(qE)$,$x'=K'/(qE)=\alpha^2K/(qE)$,$x'=\alpha^2x$;由此可得 $t_r=4x/V_{ix}$,$t'_r=4\alpha^2x/\alpha V_{ix}=4\alpha x/V_{ix,}=\alpha t_r$。从以上分析可知这两个离子在飞行时间质谱

的漂移区和反射器中飞行时间的和分别为 $t_{总} = t + t_r$，$t'_{总} = t/\alpha + \alpha t_r$，从式中可见，如 $\alpha > 1$，即具有较高能量的离子在无场区飞行时间较短，而在反射器中飞行时间较长；相反，若 $\alpha < 1$，则具有较低能量的离子在无场区飞行时间较长，而在反射器中飞行时间较短。所以这两个相同质量，初始动能不同的离子在反射器的作用下，飞行时间相互补偿，最终使这两个离子几乎同时到达检测器，提高了仪器的分辨率。反射型飞行质谱不仅提高了仪器的分辨率，而且能够测定源后衰变离子(post source decay，PSD)，这些离子所包含的信息是结构解析中的重要依据。PSD 的基本原理如图 9.18 所示。当离子被加速出离子源时，离子门开(on)时，离子 m/z 1 281.7 被消去，只剩下 m/z 1 270.7(见图 9.19)。选

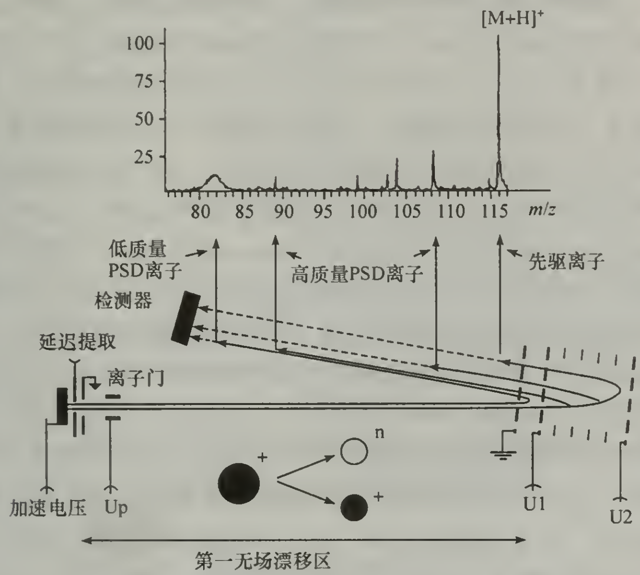


图 9.18　PSD 原理图

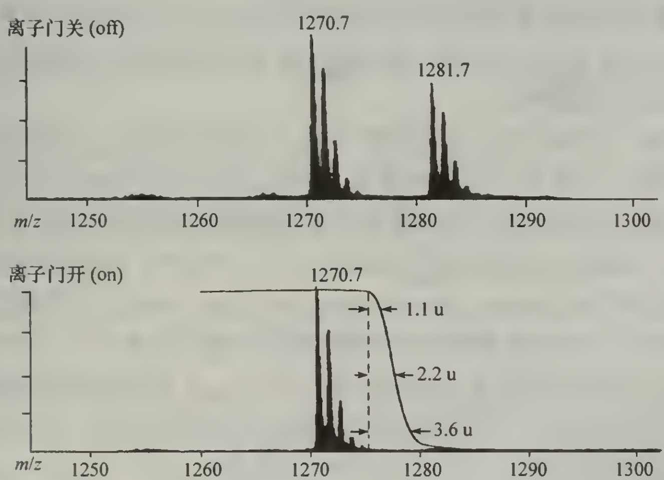


图 9.19　离子门的作用原理

择的离子在进入无场漂移区后它从离子源内解析电离过程中获得的内能，在无场漂移区内飞行过程中通过发生断裂而释放其能量。这个过程称为源后衰变(post source decay，PSD)，由此而得到的碎片离子由于它与其先驱离子具有相同的飞行速度但能量不同，所以，在反射器内穿越的深度不一可被反射器按其质量大小(即能量大小)从小到大依次反射，并到达检测器检测。在上述分析过程中，反射器的电场是一个固定的电压，实际上那些能量小的碎片离子(即低质量的离子)不能穿入反射器而无法得到检测，一般能量大于50%先驱离子能量的离子才能被检测到，为此，在反射型飞行时间质谱中反射器上的电位是逐级降低的，这样在每一级电位下可得到与其相应的较低能量的碎片离子。全部碎片离子即可采用若干级降低电位的方法得到。如在某些商品仪器中，其加速电压为28.5kV，反射器最后透镜电压为30kV。在实施PSD离子检测时，第一级采用30kV，接着为28.5，27.08，23.29，19.90……3.66kV，共分13级，由此而得到的14张质谱图可经计算机软件处理后成为一张完整的质谱图。在数据处理中，以相邻两张质谱图中同时出现的离子的峰强作归一化，所以全谱具有相同的强度标尺。在飞行时间质谱中，另一个重要的技术措施称为延迟引出(delay extraction)。它是基于Wiley和Mclaren提出的时间延迟聚焦(time lag energy focusing)概念的基础上由Brown和Lennon[28]等首先实现的。如在基质辅助激光电离中，当激光束照射样品和基质的共晶体时，样品分子获得了能量进行电离并溢出表面，此时不立即加上加速电压将离子推进漂移区，而是在经过一段时间的延迟后(几百纳秒)再加速出去，在延迟时间内离子处在推斥极及引出极之间的无场区内，由于样品分子电离后经延迟再加速所以称为延迟引出。这种技术的应用可以补偿离子初始动能分布对分辨率的影响，也就是说，初始动能不一样但质荷比相同的离子能同时到达检测器，从而提高分辨率。延迟时间的长短与测试样品性质和应用的基质有关，Karas[29]用芥子酸(sinapinic acid)和二羟基苯甲酸(DHB)作为基质对碳酸酐酶(carbonic anhydrase，分子质量为29024.6Da)进行测定，发现以芥子酸作基质时，最佳延迟时间为300ns，要比用DHB(250ns)作基质要长。即使如此，前者分辨率为700(FWHM)，而后者可达900(FWHM)，如图9.20所示。

总之，在使用基质辅助激光电离-飞行时间质谱仪进行生物大分子测试时，除了仪器本身性能之外，还要考虑样品的性质、基质的选用、样品制备的方法、激光强度等因素对测试时的分辨率、灵敏度的影响。

从以上介绍的飞行时间质谱工作原理可知，飞行时间质谱工作过程是脉冲式的。即样品分子在离子源内电离成为离子并滞留一段时间后由加速电压把它们引入无场漂移区，按它们的质量大小不同，由小到大逐个到达检测器而被检测，所以飞行时间质谱与脉冲式的电离技术，如激光电离(laser desorption，LD)、等离子体电离(plasma desorption，PD)、基质辅助激光电离(matrix assistant laser desorption，MALDI)能很好地匹配。为了实现飞行时间质谱与连续电离源如电子喷雾电离相连接，在离子引入飞行时间质谱的方法上相应发展了许多相关的技术，如在线门(in-line gate)、垂直加速器和离子阱储存(ion trap storage)等离子引入技术，以下对垂直加速器方法作简单介绍。垂直加速器的结构如图9.21所示，低能离子束从离子源出来经静电透镜聚焦后送入加速区，加速区的宽度约为漂移区的5%～10%，当离子充满加速区后，漂移区内的电极(一个或多个电极)上加上脉冲电压，建立一个脉冲的引出场，此场与连续离子流成绝对垂直，离子在引出场的作用

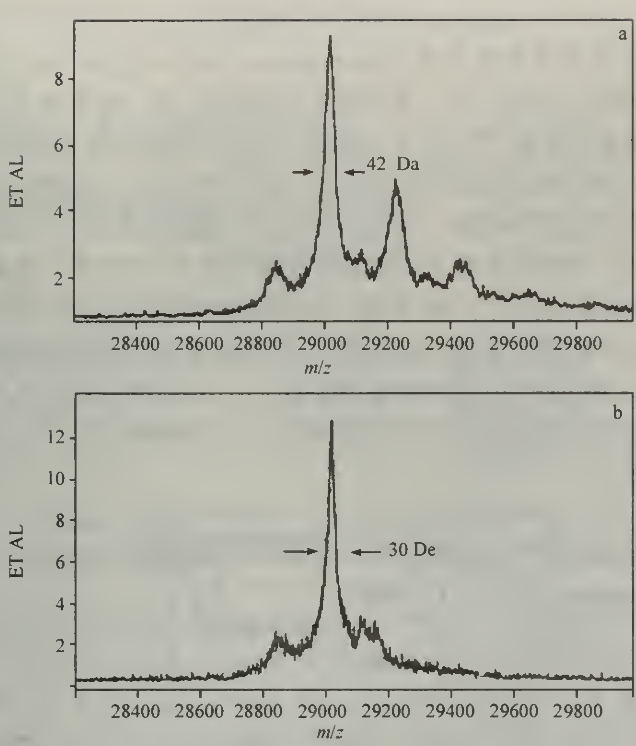


图 9.20　不同的基质和延迟时间对分辨率的影响

a. 芥子酸为基质；b. 二羟基苯甲酸为基质

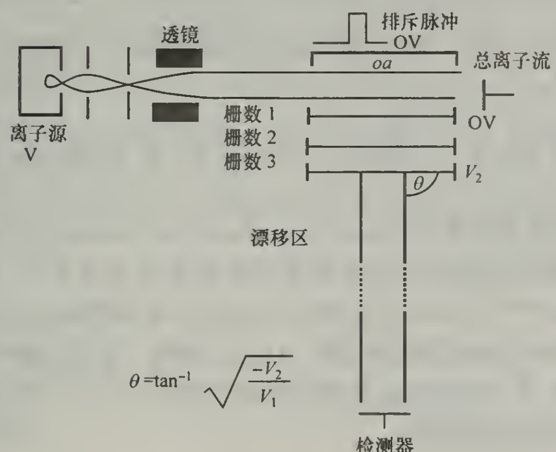


图 9.21　垂直加速器-TOF 示意图

下获得能量并飞进无场漂移区。当第一批离子离开引出场后，关闭引出场电压，使离子束再次充满加速区，如此反复循环，可实现连续离子源与飞行时间质谱相结合。由于连续离子流的初始速度方向和引出场相垂直还带来了另一个优点即离子在漂移区方向上的速度分量极小，因此可以提高飞行时间质谱的分辨率。目前采用垂直加速器离子引入技术后，其分辨率可达 5 000（FWHM），仪器已商品化。在解决了连续离子流和飞行时间质谱相

匹配的问题后，为了在 TOF 仪器上实现 MS/MS 功能，Morris 于 1996[15]年发表了他设计的四极-垂直飞行时间质谱串联装置（quadrupole-oa time of flight mass spectrometer，Q-TOF），具体结构如图 9.22 所示，第一级质谱采用四极杆，第二级质谱为反射式飞行时间质谱，此两级质谱成垂直配置，其间还有碰撞活化室以实现碰撞诱导解离（CID）。进行质谱-质谱实验时，第一级四极杆质谱用于选取单一离子并将它送入碰撞活化室与惰性气体发生碰撞并使母离子发生诱导裂解。碰撞活化室由六级杆组成，在工作状态下极杆上仅有射频电位（RF-only），因而所有离子均能通过碰撞活化室，到达第二级垂直飞行时间质谱的加速器中，在推斥极的作用下，离子进入 TOFMS 进行质量分离。由于此仪器配以电子喷雾电离源（ESI）和基质辅助激光电离源，可分别实现与高效液相色谱联用或直接进样。目前已成为蛋白质组研究工作中的重要手段之一，在商品仪器中已可见采用多级反射器（W 型）的飞行时间质谱作为第二级质谱仪。由于多级反射，其分辨率更高，可达 20 000（FWHM）左右。

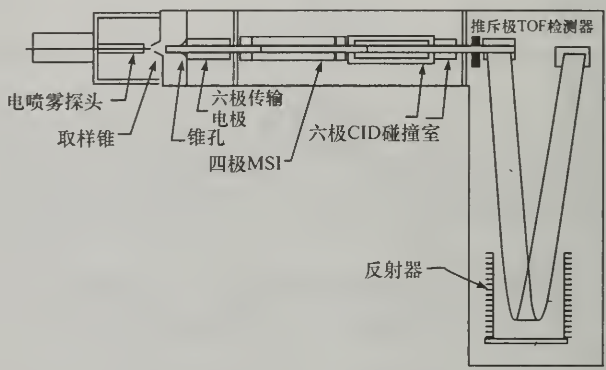


图 9.22　Q-TOF 结构示意图

第四节　傅里叶转换回旋共振质谱

傅里叶转换回旋共振质谱仪（Fourier transform-ion cyclotron resonance mass spectrometry，FTMS）由于其超高分辨率及质量测量的准确度而备受关注。第一台 FTMS 是由 Comisarow 和 Marshall[30]于 1974 年研制的，但是它在蛋白质和多肽分析中的应用还是在电子喷雾电离（ESI）和基质辅助激光电离（MALDI）的引入后才逐渐广泛起来。FTMS 主要有四大部分所组成，即：超导磁场（其磁场强度可高达 3～9T），分析池（analyzer cell），超高真空系统和计算机数据处理系统。

一、基本原理

FTMS 的工作原理是基于离子在均匀磁场中的回旋共振运动。当离子在进入分析池后即被“陷入”池中，由于分析池置于均匀的强磁场中（B），此离子将受到洛仑兹力的作用，$F_L = evB$，式中 e 为离子的电荷，v 为离子的速度，B 为磁场强度。由于离子在磁场中受力方向始终保持与速度方向垂直，所以离子在磁场中将作圆周运动，此时离子受到的洛

仑兹力和离心力相平衡,见图 9.23。$F_L = F_Z$,即

$$evB = \frac{mv^2}{r} \tag{9-4-1}$$

式中 m 为离子质量,r 为离子回旋半径,因为 $v/r = \omega = 2\pi f$,所以

$$eB = m\omega = 2m\pi f \tag{9-4-2}$$

式中 ω 为角速度,f 为回旋的频率(Hz),由此可得

$$f = eB/2\pi m \tag{9-4-3}$$

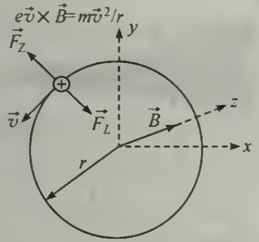


图 9.23　离子在磁场中的受力图

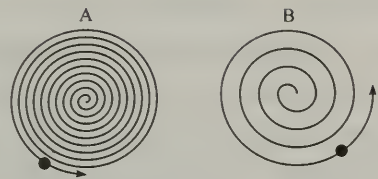


图 9.24　不同质荷比离子的回旋频率图

可见离子的回旋频率与离子的质荷比成反比,而与磁场强度成正比。图 9.24 显示了两个不同质量的离子在磁场中回旋共振时的情况,离子 A 由于质荷比比 B 小,所以其回旋的频率比 B 高。从(9-4-2)式又可见对于一个给定质量的离子,若置于较高的磁场强度内,其回旋频率亦较高。由于在高频信号测量中,较高的频率可得到较好的信噪比,所以采用高磁场能得到较好的测量结果。另外,从(9-4-2)式也可推论,离子的回旋频率与它的初始速度无关,所以离子的动能分布不会影响 FTMS 的性能,这就是 FTMS 可以得到超高分辨的基本原因之一。分析池是 FTMS 中的关键部件,其结构见图 9.25,它是一个立方体,它的六个面构成了相互绝缘的三组极板,上下一组称为接受极(receiver plate),用于接受离子产生的信号。左右一组为传送极(transmitter plate),用于传送激发脉冲信号。前后一对为“陷电极”(trapping plate),它与磁场相垂直,为了阻止离子沿磁场方向从池中逸出,通常在此组极板加上一个正或负 1～2V 的电压,当离子进入分析池后,即使质量相同的离子,由于其初速度不同,在磁场中做回旋运动时,其回旋频率相同而半径不同。

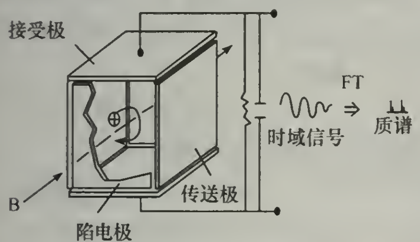


图 9.25　分析池示意图

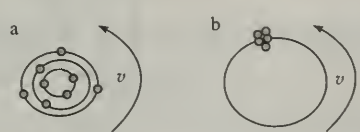


图 9.26　不同初速度的离子的回旋状态图

图 9.26a 显示了这些离子在不同的半径上作回旋运动的情况,如果此时在传送极加上一个频率与这些离子回旋频率相同的激发脉冲,这些离子将与其产生共振,离子在获得了能量后回旋半径逐渐增大,最后趋向一致。所以,这些离子在激发后回旋在相同的半径轨道上并聚集在一起。如图 9.26b 所示,轨道半径由下式表示 $r = \mathrm{Arf}\times t/B$,式中 Arf 是

激发脉冲的大小，t 为激发脉冲的时间。从上式可见，轨道半径与激发脉冲的大小、时间成正比，与磁场强度成反比。FTMS 中离子信号的接受是在接受极板上完成的，其基本原理见图 9.27，当正离子在池中做回旋运动并接近接受极时，它吸引了与接受极相连的外电路中的电子，感应出"像电流"(image current)。所以，在离子的激发过程中，若在传送极上加上很快的射频扫描(RF chirp)，其频率范围能覆盖欲测定的质量范围所相应的频率，池中的所有离子将同时被激发，并在接受极上获得与其相应的像电流，如图 9.28 所

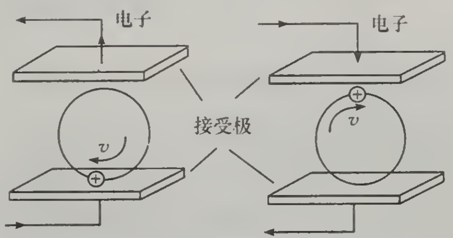


图 9.27　像电流产生示意图

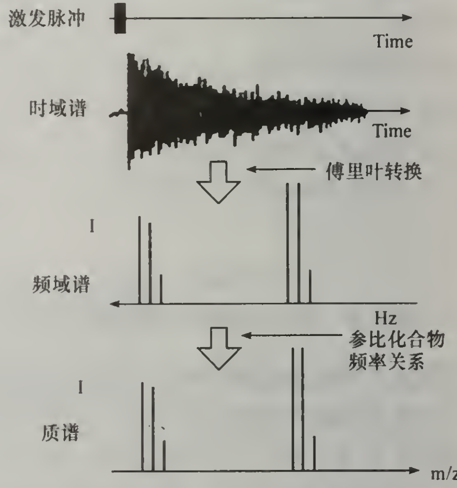


图 9.28　FTMS 转换流程图

示。这种图谱称为时域谱(time-domain spectrum)，为了从时域谱中得到每一个质荷比的离子的频率成分，通常用傅里叶转换方法来进行计算，经过这种转换的谱称为频域谱(frequency-domain spectrum)。若以参比化合物建立了频率和质量关系，频域谱即可转化为质谱。为了实现连续测试，池中剩余离子必须在每次测试后从池内清除(qunch)，它是通过分别加在一对陷阱极上的正负电压来实现的。使正离子(负离子)与负极(正极)相碰撞，这样剩余的离子即可被消去。综上所述，FTMS 的操作是通过四个步骤，即清除、电离、激发和检测程序来完成的。如图 9.29a 所示，首先清除池内的剩余离子，接着用各种离子引入技术如四极杆或静电透镜组将电离后的离子引入分析池中，然后以频扫(RF

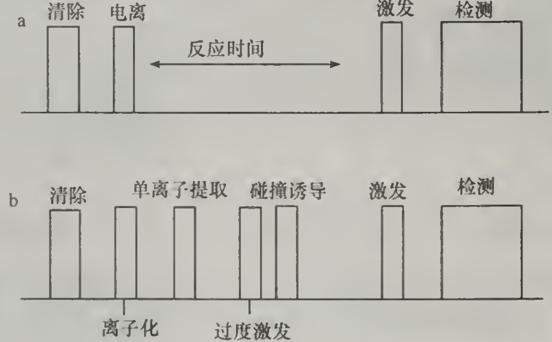


图 9.29　FTMS 脉冲序列图

chirp)、脉冲激发(impulse excitation)、储存波形逆向傅里叶转换(storage wave inverse Fourier transform)等方法同时激发不同质荷比的离子,最后同时记录所有不同回旋频率的信号。

二、质量分辨率、质量精度和质量范围

(1)质量分辨率

在 FTMS 中质量分辨率可用下式表示 $R=\frac{1}{2}fT$,式中 f 为回旋频率,T 为采集信号持续时间,从上式可见,FTMS 的质量分辨率是与磁场强度(B)和信号持续时间(T)成正比,而与离子的 m/z 成反比。为了达到高分辨,在磁场强度确定后,分辨率就取决于信号持续时间(T)的长短,虽然 FTMS 的分析池是处于高真空状态下,但是离子和池中的中性分子发生碰撞仍旧是不可避免的,这种碰撞会减低离子的速度以至于逐渐减少离子回旋的半径甚至使信号消失,所以在 FTMS 仪器中,保持超高真空是必需的,通常其真空度应优于 10^{-9}torr,这样才能保证较长信号时间,以得到高分辨数据。同时我们也应看到基于 FTMS 原理,它的分辨率会随着所测的分子质量的增大而降低$\left(f=\frac{eB}{2\pi m}\right)$。Mclaffety[31]曾以细胞色素 c 为标样对 FTMS 性能进行测试,其结果见图 9.30,图中显示,随着质量测量范围的增大,分辨率呈下降趋势。尽管如此,FTMS 的分辨率还比其他类型的质谱高。

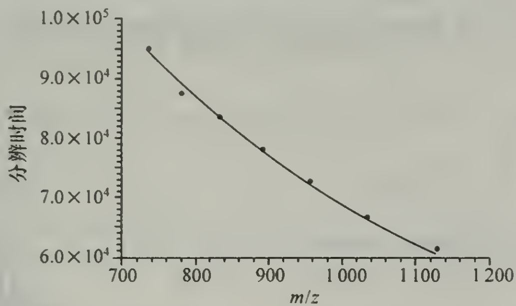


图 9.30 质量分辨率与质量范围关系图

(2) 质量准确度

FTMS 的高分辨性能为质量测定的准确度提供了前提,但并不意味着分辨率越高,质量测量准确度也越高。在良好的质量标尺校正情况下,即使测量离子的 $m/z>1\ 000$,FTMS 的质量测量准确度仍可达到 1～10ppm。这对于肽谱检索的准确度是非常重要的。

(3) 质量范围和灵敏度

对于生物大分子的分析,质谱仪的质量测定范围亦是选用质谱仪的主要指标之一,FTMS 理论可测质量的上限与磁场强度(B)和陷阱板上的电位(V_T)有关,磁场强度越高,可测的质量范围越大,但是陷阱极上的电位所形成的径向电场会引起离子从径向排出池外,故而陷阱电位应尽可能低,如图 9.31 所示,在 $V_T=1$V 时,磁场强度为 7.5T 的可测极限为 280 000(m/z),而 3.0T 只能达到 50 000(m/z);同样,对某一强度的磁场如

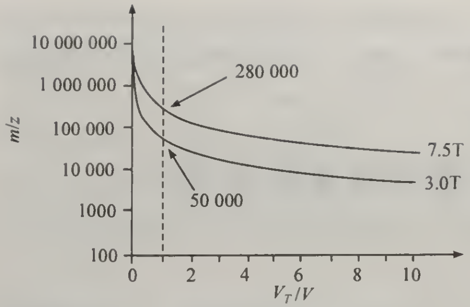


图 9.31 质量范围与陷阱电极上电位 V_T 的关系

7.5T 随着 V_T 的增加其可测定质量范围呈下降趋势,大体上它们成线性关系,FTMS 的灵敏度随着磁场强度的增强而增高,基本上也呈线性关系,同时延长采集数据时间并叠加采样信号均对提高灵敏度和信噪比有益。

(4) 多级质谱-质谱功能

实现质谱-质谱是 FTMS 的重要功能之一。在常规测试的脉冲序中,插入 3 个脉冲,即单离子提取,过渡激发,碰撞活化,即可实现,如图 9.29b 所示。在清除电离脉冲后,以一定形式的激发脉冲,消去除选择的先驱离子以外的所有离子,然后以过渡激发脉冲激发先驱离子使其回旋半径增大并聚集在一起,接着向池内注入一定量的气体,使离子与气体分子发生碰撞并发生裂解,最后再激发所有的碎片离子并接受其信号,反复整个上述过程,即可实现多级 MS/MS 功能,此功能只有在 FTMS 和离子阱质谱中才能实现。另外,由于电喷雾电离形成的蛋白质多电荷离子因其电荷间的相互排斥作用,故而在碰撞诱导解离过程中比单电荷离子更容易发生裂解并再生成不同电荷状态的离子。这些离子的准分子离子与其同位素峰在 FTMS 实验中由于它的超高分辨能力很容易得到分离,我们只需观察在一个质量单位内所能包含的峰数即为离子所带的电荷数。图 9.32 所示为某一带有 14 个电荷的肽段的质谱图。可见峰间质量差仅为 0.07。由于 FTMS 的超高分辨性能和多极 MS/MS 功能,并实现了与 MALDI 电离源和 ESI 电离源相连接,所以它在蛋白质结构分析中逐渐地被广泛应用,惟一的缺点就是价格昂贵。

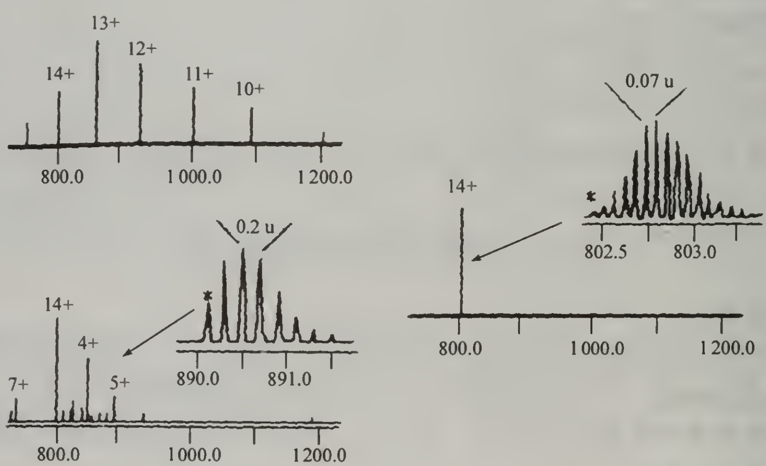


图 9.32 带 14 个电荷的离子的同位素分布图

(杨一鸣 孙海芳)

参考文献

1 Karas M, Hillenkamp F. Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. Anal Chem, 1988,60:2299～2301

2 Fenn JB, Mann M, Meng CK *et al*. Electrospray ionization for mass spectrometry of large biomolecules. Science, 1989, 246:64～71

3 Loo JA. Bioanalytical mass spectrometry:many flavors to choose. Bioconjugate Chem, 1995,6:644～650

4 Loo JA. Studying noncovalent protein complexs by electrospray ionization mass spectrometry. Mass spectrum Rev, 1997,16:1～23

5 Wood TD, Chorush RA, Wampler FM 3rd *et al*. Gas-phase folding and unfolding of cytochrome cations. Proc Natl Acad Sci U S A, 1995,92:2451～2454

6 Young MM, Tang N, Hempel JC *et al*. High throughput protein fold identification by using experimental constraints derived from intra molecular cross-links and mass spectrometry. Proc Natl Acad Sci U S A, 2000,97:5802～5806

7 Rappsilber J, Siniossoglon S, Hunt EC *et al*. Generic strategy to analyze the spatial organization of multi-protein complexes by cross-linking and mass spectrometry. AnalChem, 2000,72:267～275

8 Yost RA, Enke CG. Triple quadrupole mass spectrometry for direct mixture analysis and elucidation Anal Chem, 1979, 51:1251A～1261A

9 Loo JA, Edmonds CG, Smith RD. Primary sequence information from intact proteins by electrospray ionization tandem mass spectrometry. Science, 1990, 248:201～204

10 Schwartz JC, Jardine I. Quadrupole ion trap mass spectrometry. Methods Enzymol, 1996,270:552～586

11 Jonscher KR, Yates JR 3rd. The quadrupole ion trap mass spectrometer—a small solution to a big challenge. Anal Biochem, 1997,244:1～15

12 Qin J, Chait BT. Identification and characterization of post translational modifications of protein by MALDI ion trap. Anal Chem, 1997,69:4002～4009

13 Qin J, Fenyo D, Zhao Y *et al*. A strategy for rapid, high-confidence protein identification Anal Chem, 1997,69: 3995～4001

14 Marshall AG, Hendrickson CL, Jackson G S. Fourier transform ion cyclotron resonance mass spectrometry: a primer. Mass Spectrom Rev, 1998,17:1～35

15 Wood TD, Guan Z, Border CL *et al*. Proc Natl Acad Sci USA, 1998,95:3362～3379

16 Morris HR, Paxton T, Dell A *et al*. High sensitivity collisionally-activated decomposition tandam mass spectrometry on a novel quadrupole/orthogonal-acceleration time-of-flight mass spectrometer. Rapid Commum Mass Spectrom, 1996, 10:889～896

17 Borchers C, Parker CE, Deterding LJ *et al*. Preliminary comparison of precursor scans and liquid chromatography-tandem mass spectrometry on a hybrid quadrupole time-of-flight mass spectrometer. J Chromatogr A, 1999,854:119～130

18 Borchers C, Peter JF, Hall MC *et al*. Identification of in-gel digested proteins by complementary peptide mass fingerprinting and tandem mass spectrometry data obtained on an electrospray ionization quadrupole time-of-flight mass spectrometer. Anal Chem, 2000, 72:1163～1168

19 Zhang Z, McElvain JS. De novo peptide sequencing by two-dimensional fragment correlation mass spectrometry. Anal Chem, 2000, 72:2337～2350

20 Randall K,Julian Jr, Cooks RG. Broad-band excitation in the quadrupole ion trap mass spectrometer using shaped pules created with inverse Fourier transform. Anal Chem, 1993,65:1827～1833

21 Soni MH, Cooks RG. Selective injection and isolation of ions in quadrupole ion trap mass spectrometry using notched waveforms created using the inverse Fourier transform Mass Spectrometry method using notch filter Anal Chem, 1994, 66:2488～2496

22 Kelly PE. Mass spectrometry method using notch filter. US Patent. 5134286 1992

23 Kelly PE. Mass spectrometry method using filtered noise signal. US Patent. 5206507 1993

24 Mcluckey SA, Georinger DE, Glish GL. Selective ion isolation/rejection over a brand mass range in the quadrupole ion trap. J Am Soc Mass Spectrom, 1991,2:11～21
25 Wiley WC, Mclaren JH. Res Sc: Instrum, 1955,26:1150～1153
26 Macfarlane BD, Skowronski RP, Torgerson DF. Biochem Biophys Res Commun, 1974,60:616～620
27 Mamyrin BA, Karatajev VJ *et al*. Soviet Phys JETP, 1973, 37:45～48
28 Brown RS, Lennon JJ. Mass resolution improvement by incorporation of pulsed ion extraction in a matrix-assisted laser desorption/ionization linear time-of-flight mass spectrometer. Anal Chem, 1995,67:1998～2003
29 Bahr U, Stahl-Zeng J, Gleitsmann E *et al*. Delayed extraction time-of-flight MALDI mass spectrometry of proteins above 25,000 Da. J Mass Spectrom, 1997, 32:1111～1116
30 Comisarrow MB, Marshall AG. Fourier transform ion cyclotron resonance spectroscopy. Chem Phys, Letter, 1974, 25:282～286
31 Beu SC, Senko MW, Quinn JP *et al*. Fourier-transform electrospray instrumentation for tandam high-resoultion mass spectrometry of large molecular. J Am Soc Mass Spectrom, 1993,4:557～565

第十章　毛细管电泳-质谱联用

第一节　介　　绍

20 世纪 90 年代以来，随着基因组计划的实施和快速发展，迄今为止已有流感嗜血杆菌、大肠杆菌、酿酒酵母等几十种微生物的全序列被测定；对 cDNA 和表达序列标记的大规模测序也取得显著进展，这将为生物学家们提供前所未有和极其丰富的遗传信息。但同时人们也越来越清楚地认识到，要单单从基因的序列推断出基因产物的行为和功能是十分困难甚至是不可能的。因为即使基因被转录，它的表达还可能在翻译水平上受到调控，并且其蛋白质表达产物还将通过翻译后修饰、改变半衰期和装配到蛋白质复合物中等方式进一步受到调控[1]。为了对生物系统的功能和调控有深入的了解，对组成和调控这样一个系统的蛋白质的分析就显得十分重要。但是，现在还没有一种常规的方法可以把代表某一过程中特殊状态形式的蛋白质或代表某一特定蛋白质修饰形式的蛋白质进行扩增。与 DNA 可以通过 PCR 方法方便地进行扩增所不同的是，在对蛋白质进行分析之前必须把它们从其生物来源中分离出来。由于许多蛋白质（尤其是特殊的调控蛋白）常常是丰度极低的，所以蛋白质分析技术的高度灵敏性就极为重要。

在常用的蛋白质和多肽分离技术中，毛细管电泳（capillary electrophoresis, CE）具有最高的灵敏度。经过适当标记的蛋白质甚至可以通过激光诱导荧光检测方法在单分子水平上得到检测[2]。有两个因素对 CE 的广泛应用形成了限制：CE 通常所用的检测器都只是检测到被分析物的存在，而不能区分所分离物质各自的性质；CE 只能容许几十纳升（nanoliter, nL, 10^{-9}L,）的进样体积，所以其浓度检测极限与层析方法相比并没有显著的优越性。但是，后来发展的在线预浓缩方法能够显著提高浓度检测极限，而且质谱（mass spectrometry, MS）及串联质谱（tandem mass spectrometry, MS/MS）的引入使得在低 fmol（femtomole, 10^{-15}mol）甚至 amol（attomole, 10^{-18}mol）水平上对经 CE 分离的多肽进行鉴定成为可能[3, 4]。正是由于毛细管电泳-质谱联用（CE/MS）技术的这些显著优点，使得它越来越受到生物学家的关注，并被广泛地用于蛋白质和蛋白质组[10]的研究之中。

一、概　　况

在 20 世纪 80 年代早期，Jorgenson 和 Lukacs 指出，CE 可以产生快速和高分辨率的分离，带电物质根据它们电泳迁移率的不同而得到分离[5]。CE 的样品操作和进样自动化更加容易；可以直接采用紫外吸收及荧光发射检测器进行在线检测；进样量和消耗溶剂少。另外，对比液相层析而言，由于不存在纵向分子扩散和质量转移，CE 的峰型比较尖锐，而且质量灵敏度也常常要高得多。正是由于 CE 的这些引人注目的优点，自 80 年代后期第一台商品化 CE 仪器出现以来，CE 在技术上和应用上得到了“爆炸”式的发展。CE 在进样方法、检测器灵敏度、毛细管表面衍生、涂层技术和引入新的电泳缓冲液系统等方

面的发展进一步推动了 CE 在多个学科领域的应用。CE 能处理和分析极小体积样品的能力为使用 CE 来挑战极端的分析任务(如对单个细胞的分析)提供了基础。另外,CE 具有多种电泳形式,如区带电泳(capillary zone electrophoresis, CZE)、等电聚焦(capillary isoelectric focusing, CIEF)、等速电泳(capillary isotachophoresis, CITP)、电动层析(capillary electrokinetic chromatography, CEC)、胶束电动层析(micellar capillary electrokinetic chromatography, MEKC)和凝胶电泳(capillary polyacrylamide gel electrophoresis, CGE)等,还具有多种控制进样条件和分离特异性的方法。因此,使用 CE 来进行样品分析灵活性极大,它被广泛地用于复杂混合物的分析,如蛋白质酶解产物、药物的底物和代谢物、DNA 加合物,以及生物工程产品等。不少文章对 CE 的最新研究和应用情况作了综述[6, 7]。

尽管 CE 的多功能性、高分辨率和高速度使它在多个领域得到了广泛的应用,当时其检测方法的局限性限制了它在这些领域研究中的进一步深入。最近 10 年来,随着质谱技术的飞速发展,电喷雾离子化(electrospray ionization, ESI)质谱[8]和基质辅助激光解吸离子化(matrix-assisted laser desorption ionization, MALDI)质谱[9]已经成为生物大分子研究中必不可少的支撑技术。把 MS 作为 CE 的检测器受到越来越多的关注。MS 可能是 CE 最强有力的检测器。因为 MS 可以提供多达几千个的离散性和选择性的“检测器”等价物(每一个质荷比值都可以看作是一个单独的检测器),它们并行地工作并能提供完整的“分子离子”的分子质量信息和母离子在质谱仪中解离而产生的结构相关信息。从某个角度来看,CE 和 MS 的联用简直可以说是理想化的——CE 的分离基于溶液中离子的不同迁移率,而 MS 则根据离子在气相中的质荷比(mass-to-charge ratio, m/z)来分析离子——这两种高度正交化的分析方法分别利用了在两种截然不同的环境中的离子迁移运动。CE 分离中常常遇到的迁移时间变化等问题可由 MS 检测的选择性和特异性得到补偿。把分离技术与质谱联用还可以显著提高检测极限并且简化对分析结果的解析过程。

Jorgenson 等人的早期工作已经清楚地揭示出 CE 的潜力,但是要把 CE 与 MS 联接起来还需要解决几个概念上和实践上的问题。首先,CE/MS 必须同时建立起 CE 的电场梯度和 ESI 的高压。其次,CE/MS 的接口必须与 CE 的低流速(通常≤1μl/min)相容。第三,要避免任何可能由毛细管两端压力差造成层流而引起的峰型变宽,或者与检测密切相关的死体积。另外,CE/MS 要有实际应用价值,其检测灵敏度还必须在亚 pmol(picomole, 10^{-12}mol)水平之上。80 年代早期就有了为 LC/MS 设计的热喷雾(thermospray)接口,但这种接口并不适合用于 CE/MS,因为它的流速要求在 100μl/min 以上,而且也不够灵敏。1984 年,Fenn 等人首次将 ESI 离子源与 MS 连接[10],其中液体从毛细管的一端以 1～10μl/min 之间的速度流出并在高压电场中被雾化而进入 MS。这个结果使 Smith 等人深受鼓舞并致力于发展基于 ESI 的 CE/MS 在线联用技术。1987 年,Smith 等[11]人发表了有关 CE/MS 联用的第一篇论文,对 CE/MS 的研究情况作了综述并客观地评估其潜力和限制因素。此外,Moseley 等[12]人还发展了基于连续流快原子轰击(continuous-flow fast atom bombardment, CF-FAB)离子化方法的 CE/MS 接口。近年来,将 CE 与 MS 联用的技术已经达到了一个相当成熟的程度,这当然是与各种 CE/MS 接口的不断完善密不可分的。

CE 已经与多种不同类型的质谱质量分析器联用,其中包括三级四极杆(triple quadrupole, TQ)质谱、傅里叶转换离子回旋共振(Fourier transform ion cyclotron reso-

nance, FTICR)质谱、离子阱(ion trap, IT)质谱和飞行时间(time-of-flight, TOF)质谱。尽管在将 CE 与 CF-FAB 和 MALDI 联用方面作了不少努力,ESI 还是人们在将 CE 与 MS 联用时的首选方法。这主要是因为 ESI 将分子从液相直接转移到气相之中,CE 与它的整合比较容易实现。而且,ESI 条件下,大分子物质是带多电荷的,这使得分子质量较高的物质可在大多数质谱仪有限的质量范围内被检测到。

对 MS 仪器的改进使它的灵敏度、速度,以及由这种强有力的检测器获得的分子结构信息的细节进一步得到提高。MS/MS 和多级质谱(MS^n)方面的研究进展更为我们展示了能提供更高选择性和结构信息的未来质谱仪器的美好前景。尽管 MS 仍是 CE 最复杂和最昂贵的检测器之一,但其费用正在降低,“用户亲切性”正在增加,而且从这种联用中可以得到的不断增加的丰富信息还将持续地减轻我们对这些“缺点”的顾虑。典型的 CE/MS系统示意图可用图 10.1 来表示。对 CE/MS 的发展和应用已有不少综述[13~17]。

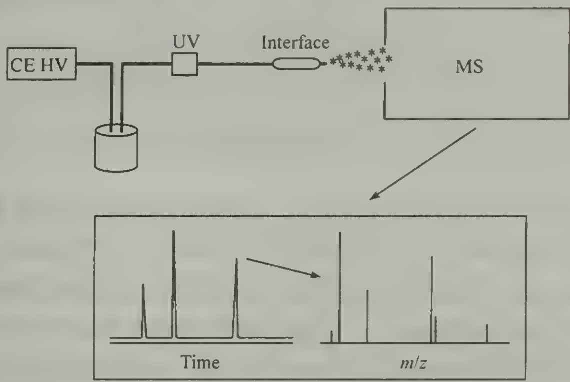


图 10.1　CE/MS 仪器设置示意图

二、CE/MS 联用的接口

从 MS 的角度来看,CE 和 MS 的联用要依赖于接口才能在不牺牲毛细管电泳分辨率的前提下在线地把分析物从电泳毛细管中的液相中高效地转移到气相中进行 MS 分析。CE 本身非常适合于分离容易在溶液中形成离子的极性化合物,但是这些类型的化合物对常规的离子化方法,如电子和化学离子化,却是一个挑战,而且不挥发性或热不稳定性化合物在此条件下容易产生热分解。直到 ESI、MALDI 和 FAB 等“软电离”方法发展起来后,这种矛盾才得以解决。CF-FAB 和 ESI 以及气动辅助 ESI 等离子化方法都已成功地与 CE 进行了联用;而等离子解吸(plasma desorption, PD)和 MALDI 等解吸附离子化方法与 CE 的联用基本上还处于非在线阶段。由于前述 ESI 方法的诸多优点,CE 与 MS 的联用多采用 ESI 方法来离子化。

最佳的 CE/MS 接口应该能提供最高的离子转移效率,最好的电导通性和喷雾稳定性,同时只引入最少的物理和化学问题。现在用来进行 CE/ESI-MS 联用的接口有三种:同轴鞘流液(coaxial sheath-flow)接口、液-液连接(liquid-junction)接口和无鞘流液(sheathless)接口(图 10.2)。

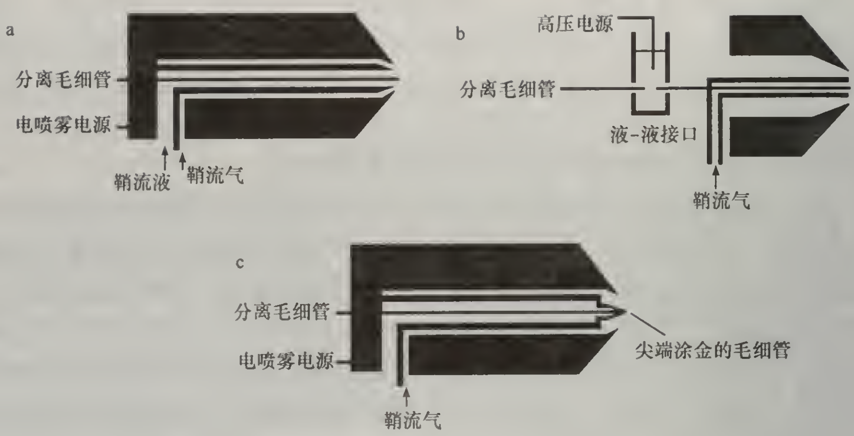


图 10.2　CE/MS 接口示意图

a. 鞘流液接口；b. 液-液连接接口；c. 无鞘流液接口

(一) 鞘流液接口

鞘流液接口(图 10.2a)首先由 Smith 等[18]人提出，它是一种最常用的 CE/MS 接口，因为它可以很容易且重复地建立起来。这种接口使用了一种同轴的层层包裹结构，分离毛细管的出口端既是 CE 的出口端也是 ESI 的发射口，其外部有鞘流液流通，以建立起 CE 分离和 ESI 发生所必需的电连接。分离毛细管出口端突出在鞘流液针管尖端之外约 1mm，CE 的流出物质与鞘流液在这里混合并被喷雾。鞘流液通常是一定比例挥发性酸与挥发性醇的水溶液(鞘流液可以综合考虑表面张力、pH 值及添加剂等因素以优化 ESI 条件)，它可以稀释 CE 的电泳缓冲液，也就减轻了可能存在的电泳缓冲液与 ESI 不相容的程度，但是在使用前必须将鞘流液脱气以防止 ESI 时气泡的产生。当液体的流速较大时，还可以在鞘流液的外部再使用鞘气来辅助喷雾。这种接口由于鞘流液的引入可以增加在选择 CE 缓冲液方面的灵活性。在分离毛细管内壁用中性化合物涂层后电渗流几乎为零的情况下，常常要使用这种类型的接口。

鞘流液中不可避免地含有离子化或中性的物质，它们在 ESI 的过程中会竞争本来就有限的电荷，从而降低了分析物的离子化效率。使用鞘流液还引入了另一些化学背景，它们可能会对分析物的检测造成干扰而且增加检测极限，还可能改变分析物的带电荷分布形态。同样值得一提的是，由于毛细管中电场梯度的存在，鞘流液和电泳缓冲液中的阴离子都朝阳极移动。当这两种溶液中的阴离子不同时，就会在 CE 毛细管的内部形成一个移动的离子边界。由于这个移动边界内部的 pH 可能会不同，当分析物进入到这个边界时，电渗流的速度、分析物的有效电荷以及它们的迁移率都会发生改变。这种现象的发生将不仅影响分析物的迁移时间和迁移顺序，而且会引起分离效率的损失。将 CE 的入口端抬高以提供除电泳流动以外的流体力学流动，可以将这种分离效率的损失保持在最小的水平。

(二) 液-液连接接口

液-液连接接口(图 10.2b)首先由 Lee 等[19]人发展起来。这种类型的接口使用了一

个 T 型结构,CE 分离毛细管和 ESI 发射毛细管之间的电连接通过由另一入口引入的缓冲液(也称构成液,make-up liquid)来建立。CE 毛细管和 ESI 发射毛细管之间的缝隙只有 10～20μm。构成液的传送速度可以通过可控流速的方式来控制,液-液连接处还可以方便地进行精确定位。

Severs 等[20]人利用透析膜建立电连接还发展了一种透析液-液连接(dialysis junction)接口。Yang 等[21]人通过这种接口进行了 CIEF/ESI-MS 实验。与常规的鞘流液接口相比,这种接口的灵敏度要高一个数量级,而且离子强度和信噪比也高许多。这种接口设计可以获得很高的分辨率,而且接口处的透析液-液连接所提供的柱后酸化过程可产生更加有效的变性作用而使蛋白质带上更多的电荷,从而提高其离子化效率。

由于液-液连接接口中 CE 毛细管与 ESI 发射毛细管是断开的,所以 ESI 发射毛细管可以很容易地被更换,与之相关的问题也就与 CE 分离毛细管隔离开来,从而延长了其使用寿命。另外,这种接口可以适应多种不同的流速条件,而且在柱后引入缓冲液使得在选择缓冲液的适当 pH 值、组成和流速以优化 ESI 效果方面具有很大的自由度。但是,CE 分离毛细管和 ESI 发射毛细管之间的不良安装将直接导致峰型变宽和分离效率的损失。液-液连接处的溶液混合或流速过低可能造成的峰型变宽和背景噪声也是这种接口的缺点所在。

(三) 无鞘流液接口

鞘流液接口和液-液连接接口的一个共同缺点是:它们的工作都依赖于含有带电荷物质的额外液流的存在,从而不可避免地在不同程度上对检测灵敏度造成影响。这促使人们致力于发展一种更加灵活、可靠而且不再需要额外液流的 CE/MS 接口。

无鞘流液接口(图 10.2c)由 Olivares 等[11]人首先提出。早期的无鞘流液接口在电泳毛细管的外部套有一根不锈钢针管,在毛细管的末端和不锈钢针管上都有一层银的沉积层。因为获得稳定的电喷雾所需的起始电压随着喷雾尖端外径的减小而降低,所以使用较小内径的毛细管可以在产生电晕放电现象之前使水溶液喷雾。这种概念被 Gale 和 Smith 用来改造早期的接口。他们不再使用不锈钢针,而是使用末端被氢氟酸蚀刻并打磨成尖端的小内径(5～10μm)熔融石英毛细管,毛细管尖端的外部用金涂层以建立电连接,并用 SF_6 作为鞘流气以抑制喷雾尖端处电晕放电现象的产生。

ESI 发射口尖端的涂层技术决定了其性能和使用寿命。Kelly 等[22]人的方法最为成功,他们将毛细管尖端先进行金的喷溅,然后再通过电镀来加强金的涂层。Chang 等[23]人报道了一种简单的方法来制备碳涂层的毛细管尖端用于无鞘流液接口。制备而成的毛细管尖端可以在比较极端的酸碱条件下使用,而且性能与金涂层的毛细管尖端相当。

无鞘流液接口可以说是把 CE 与纳升电喷雾(nanospray, nESI)或微升电喷雾(microspray,μESI)质谱联用的最简单的一种接口,而且它与这些微量喷雾技术较低的流速非常相容。这种接口喷雾电压更低,喷雾口与 MS 入口之间距离更短,这些特点使得它的质量转移(mass transfer)效率和液体微珠的去溶剂化效率大为提高。因此,使用这种接口可以获得很高的离子化效率和较低的质量检测极限,它是最灵敏的一种 CE/ESI-MS 接口。另外,与鞘流液接口相比,没有额外的化学物质被引入。

使用这种接口时,毛细管尖端的内径大小应该仔细考虑,以保持合理的离子迁移时间

和 ESI 的稳定性。例如，若是毛细管尖端的内径太小，将使得分析时间大大加长而且容易造成堵塞；毛细管尖端的内径太大，虽然可以产生更大的电渗流，却会牺牲 ESI 的稳定性。另外，由于毛细管经过了好几步的处理步骤，如牵拉出尖端、在毛细管尖端上进行金涂层、在毛细管内壁上进行化学涂层等，它的使用寿命是有限的。假如这些处理步骤的其中一步没有处理得当（如堵塞了毛细管尖端），那就不得不抛弃整根毛细管。对此，有人提出了另一种设计：将 CE 和 ESI 的电路整合到一个环路中，这样就可以直接在 CE 的毛细管入口端加上高压，而无需再在毛细管出口端进行金涂层了。值得注意的是，无鞘流液接口要求分离毛细管内部的液体流动比较稳定，这样也就把使用中性涂层毛细管（其中几乎没有电渗流，用以减小毛细管内壁对分析物的吸附作用）的可能性排除在外，从而限制了它在某些方面的应用。

三、其他联用方式

人们在将各种 CE 技术与不同的 MS 技术进行联用方面作了许多努力，可以说各种可能的联用方式都已经有了相关的报道。CZE、CITP、CIEF、CEC、CGE 等 CE 技术常常与 ESI 和 MALDI 类型的 MS 技术联用。以下对 CE 与 MALDI 的联用和模仿 CE 原理的微型预制设备(microfabricated device, MFD)与 MS 的联用情况作一介绍。

（一）CE 与 MALDI-TOF-MS 的联用

MALDI-TOF-MS 在生物大分子的分析中应用极为广泛，尤其适合于分析分子质量较大的分子和杂质较多的样品，其特点是一般只产生带单电荷或双电荷的离子，图谱的解析比较方便直观。由于 MALDI 是将离子从固相中转移到气相的一个过程，而 CE 则是一个液相分离过程，这个矛盾使得 CE 与 MALDI-TOF-MS 的联用基本上还处于非在线状态。通常，通过 CE 分离后的物质在它们流出毛细管时被收集在装有溶剂的小型收集管中。然后，再将这些收集的级分(fraction)与基质混合并点样在探测器(probe)上，用 MALDI-TOF-MS 检测。为了简化样品的处理过程，也可通过吸附作用直接将样品沉积到膜上或探测器上。此时，这些支撑物（膜或探测器）都预先用基质处理过，样品在 CE 分离完成后无需进一步处理而直接引入 MALDI-TOF-MS 进行检测。

Foret 等[24]人设计了一种巧妙的自动化 CE 级分收集接口，可以在不中断 CIEF 的情况下将样品收集在一根根短的毛细管中，然后将毛细管中的级分与基质混合后点样到探测器上进行 MALDI-TOF-MS 分析。这种接口在 CE 毛细管出口处利用鞘流液将样品条带转移到收集毛细管中。鞘流液可以在用压力移动条带的同时维持电场的存在，从而避免了条带的扩散。在非常接近接口的地方还有一个 UV 检测器可以监测样品条带的流出情况。

CE 缓冲液中的电解质与 MALDI 的相容性需要引起特别的重视。许多 CE 中常用的电解质（特别是磷酸钠盐）会抑制 MALDI 信号。Walker 等[25]人发现，在将 CE 级分直接点样到探测器上时，含有高浓度酸的基质更能耐受常常那些导致信号抑制的电解质。对经 CE 分离的多肽进行 MALDI-TOF-MS 非在线检测所达到的最高灵敏度是 Zhang 等[26]人通过膜上沉积技术获得的。用这种方法可以检测到 amol 水平的神经加压素。起始量

只有 3 fmol 的脱辅基肌红蛋白的胰蛋白酶酶解产物也通过这种方法得到鉴定。其中 CE 分离中所用的样品浓度在 pmol/μl 水平。

这种联用技术具有接近分析和鉴定低丰度蛋白质所必需的灵敏度。但是,其 pmol/μl 水平的浓度检测极限使得其应用还仅限于分析丰度较高的蛋白质。所以,CE 的在线预浓缩技术与 CE/MALDI-TOF-MS 联用技术的结合将进一步拓宽这种联用技术的应用范围。

(二) MFD 与 MS 的联用

高灵敏度鉴定蛋白质工作中的一个主要挑战就是发展能够相当或超过最灵敏分析仪器灵敏度的样品处理程序。因为现在这种工作中最常见的一类问题就是:由于吸附到塑料管内壁或其他表面而造成的样品损失;由于多步的样品处理步骤所造成的样品损失;微量的样品中污染有实验室常见的蛋白质或化合物(如人的角蛋白),等等。有人提出,建造微型预制整合型的分析模块可能是解决这类问题的一种有效方式。而且,把这种微型预制设备与质谱仪实现在线联用是一种极具吸引力的设置方式,因为 MFD 对分析步骤的自动化十分关键而且可以节约费用。MFD 的研究虽然起步较晚(1997 年),但发展十分迅速,迄今已有许多关于多种 MFD 的研究进展及其应用的报道[27~30]。这些 MFD 的共同点是:贮液池和管道都蚀刻在一块玻璃板上,另一块玻璃盖板黏接在设备的顶端,盖板上有与每一个贮液池相应的小孔。不同 MFD 的差异点在于:它们与 ESI-MS 的连接方式、实验的目的以及溶剂加入到这些设备中和从这些设备中移走溶剂的方式。最近,Wen 等[28]人建立了一种微型 CIEF 设备可与 ESI-MS 直接联用分析复杂蛋白质混合物。Xiang 等[29]人发展的 MFD 还可以通过其双透析功能净化样品。Li 等[30]人将他们的“CE 芯片”与 nESI-MS 联用。由于使用了管道的表面涂层技术、低死体积的连接方法和金涂层的 nESI 尖端,这种系统在以选择性离子监测(selected ion monitoring, SIM)模式检测时可以获得高达 3.2nmol/L 的浓度检测极限。与原来的 nESI-MS 不能进行任何预分离相比,它可以使得多组分样品在进行 nESI 的同时预先得到部分的分离。这种系统的另一个特点是它的样品分离所需要的时间极短,一般小于 1.5min,单个峰宽只有 2s 左右,需要与速度较快的 TOF-MS 联用。

尽管将 MFD 用做 ESI-MS 的前端设备的研究还处于起步阶段,但是,将几步按顺序进行的操作整合到微型设备中,以及实现无需人的干预而自动进行多步骤分析的可能性很明显是非常诱人的。这种技术必将在未来的蛋白质生化分析中扮演越来越重要的角色。

四、CE/MS 的限制性因素及其改进方法

(一) 限制因素

CE 的一个主要限制性因素是:在不降低分离效率的前提下它所能分析的样品体积极为有限。所以,CE 的浓度检测极限常常要比层析方法高几个数量级。要想使 CE/MS 成为微量生物分析的主流技术,就必须使它能够与分析物的生物来源有效而相容性地结合起来。现在生物学中绝大多数的样品制备技术都是通过手工操作的,而且很难产生体积

在微升以下的样品。所以,通常微量的样品仅以较低的浓度存在于较大的体积中。能够在不影响分离的条件下被 CE 分析的样品通常在样品总量的 1% 以下。即使在通过 MS 获得相对较高的检测灵敏度的情况下,对这种样品进行成功的分析也还要求其浓度在中等 fmol/μl 或 pmol/μl 水平。这就意味着若是没有样品的预浓缩过程,CE/MS 只能用于分析生物样品中那些主要的蛋白质。

CE/MS 的另一个缺点是它的迁移时间容易随着环境温度的变化而波动。尽管许多 CE 仪器都有温度控制设备,但还是有一大部分 CE 毛细管伸出在 CE 和 MS 仪器之间而得不到温控。对于那些标定类型的工作或其中含有未知成分的混合物的分析,适当地加入一些内标可能会有所帮助。CE 分离毛细管的内壁化学条件也对 CE 的分离有重要影响。

像 LC/MS 一样,在 CE/MS 中也要避免使用非挥发性的缓冲液成分。所以在CE/MS 中常常要在选择合适的操作条件之间作出妥协。由于使用非挥发性盐所造成的灵敏度限制和离子源堵塞现象阻碍了把 CE 分离条件向在线 CE/MS 联用的直接转移。非挥发性的添加物,如环糊精等,在进行性质相近的物质甚至光学异构体的分离时常常被用到。但是,这些添加物的浓度却常常受到应用上的限制。另外,含有表面活性剂的离子流进入到 MS 离子源的时候,会严重影响 MS 的灵敏度。

(二) 分析方法的改进

1. 通过 CITP 进行在线预浓缩

有多种技术已经被用来提高 CE 的浓度检测灵敏度。其中,通过 CITP 的在线样品预浓缩技术受到了最多的关注。与 CZE 具有很高的分离效率而其浓度检测灵敏度却不高相比,CITP 能够提供很好的浓度检测灵敏度,但对鉴定复杂混合物却缺乏足够的分辨率。将 CZE 和 CITP 结合起来就可以克服它们的缺点而保留各自的优点。有两种基本的方法可以将它们有效地结合起来:柱上短暂等速电泳联合区带电泳(transient isotachophoretic CZE, tITP-CZE)[31]和等速电泳与区带电泳联用(CITP-CZE)[32]。

CITP-CZE 系统利用了两根分离毛细管,进行 CITP 预浓缩的毛细管在前,聚焦好的样品条带进而被转移到后面的另一根毛细管中进行 CZE 分离。两根毛细管之间的连接是将 CZE 分离毛细管通过隔膜插入 CITP 毛细管中来完成。tITP-CZE 的方法是在同一根毛细管上先进行 ITP,然后再进行 CZE 分离。这种方法在易用性和与 MS 连接方面显得效率更高,它可以用于任何 CE/MS 系统中而无需做任何仪器上的改动。尽管所有的在线 CITP 预浓缩方法都会降低分离效率,但是这种方法的综合效率还是要比常规的 LC 方法高得多。

类似的还有利用 CZE 中的电场放大作用(field amplification)[33]的样品堆栈技术。

2. 通过层析方法进行在线预浓缩

CE 样品在线预浓缩的层析方法是在 CZE 毛细管中插入一个样品抽提装置,也称为膜预浓缩(membrane preconcentration, mPC)装置,它们通常由 C18 材料充填的微型柱床或含有 C18 材料的膜制成。样品用压力推入该系统中,其中的分析物被抽提出来并且聚集在微型装置上,然后再把它们洗出到 CZE 毛细管中,其体积只有几纳升。由于样品在

抽提过程中相当于被“固定化”在抽提装置上，所以，其中的分析物可以被有效地清洗以去除盐分和杂质或进行缓冲液更换。用这种方法可以达到 amol/μl 的浓度检测极限。

关于这种在线样品预浓缩技术的重要性在许多对痕量生物样品的分析中都有报道。Naylor 等[34]人进一步将这种 mPC-CE/μESI-MS 系统与 tITP 方法结合起来。样品从抽提装置中洗脱后，夹在先导缓冲液和终结缓冲液之间先进行 tITP 聚焦。浓缩好的样品随即在同一根毛细管中进行 CZE/μESI-MS/MS 分析。与普通的 mPC-CE/ESI-MS 系统相比，这种方法可将灵敏度再提高 100 倍。Figeys 等[35]人用由 C18 微珠充填的微型装置通过固相微抽提(solid-phase microextraction, SPM)浓缩作用对经过高分辨率双向凝胶电泳(two-dimensional gel electrophoresis, 2-DE)分离的酵母蛋白质进行了分析。在凝胶中 fmol 水平的蛋白质也可以得到成功的鉴定，使得 SPM-CZE/μESI-MS 成为通过 MS/MS 来进行蛋白质鉴定和序列分析的最灵敏的系统之一。

3. 使用内径较小的毛细管

Wahl 等[36]人报道，使用小内径毛细管来进行 CE 可以获得更高的灵敏度。一般来说，ESI/MS 的最佳灵敏度应该在 CE 电流大致等于或小于 ESI 电流时获得，此时 CE 往 ESI 源输送带电物质的速率最小，离子化效率最高，ESI-MS 将以一种质量敏感的检测器系统起作用。

在涂层的毛细管中进行蛋白质混合物分离时，使用 5μm 内径的毛细管检测灵敏度要比使用 50μm 的毛细管高 25～50 倍。绝对检出量达到 600 amol 水平。但在使用小内径毛细管时，样品和缓冲液的制备和处理都需要采用改良的程序，以防止毛细管的堵塞。

4. 使用减小流出速度进行检测的方法

MS/MS 仪器的工作周期限制了其检测的动态范围(即对含有不同丰度多肽的复杂混合物的分析和覆盖范围)。通常，MS/MS 仪器都被设计为按照强度降低的顺序来分离出多肽离子进行碰撞诱导解析(collision induced dissociation, CID)。在用 MS/MS 进行蛋白质鉴定时，由于每一个分析周期包括多肽质量分析、多肽离子选择、CID，以及获取所选择离子所产生的 MS/MS 质谱图，整个过程需要 1～3s。所以，在一个典型的层析或电泳峰中，只有少数几个强度最高的离子被分析到，而强度较低的离子则被忽略了。

在 CE/MS 中，“降低流出速度(reduced elution speed, RES)检测”的概念首先由 Goodlett 等人提出，它可以增加 MS 分析多肽的数目。最近，这种方法又被应用到增加 MS/MS 分析多肽的动态范围。这种被称为峰停留(peak parking)的方法通过降低从多肽分离系统向 MS 的样品流速来延长对特殊分析物的分析时间。在第一个要分析的物质进入到 ESI 源之前，电泳电压降低，于是溶液的流出速度也随之降低。这样，在单位时间内进入 ESI 源的溶液量超过它的离子化能力时，可以有更多的扫描图谱被记录下来而不会有明显的离子强度损失。总的结果就是可以使得更多的分析物离子被 MS 分析。这种将分析物进入 MS 的时间延长的方法可以用于以下方面：①增加扫描的 m/z 范围；②增加某一特定离子在其迁移过程中被 MS 扫描记录的数目；③增强对特定离子的检测灵敏度。这都对增加 MS/MS 仪器的动态范围有直接的贡献。另外，这种方法并不需要增加样品的消耗量而且几乎不会造成离子强度的损失，后者对于 MS/MS 仪器的多肽测序能力来

说极为重要。

Figeys 等[37]人进一步发展了这种技术，他们的数据依赖调节型 SPM-CZE/μESI-MS/MS 系统可以自动在各组分从 CE 系统流出时降低 CZE 的分离电压，从而实现自动化的峰停留。应用这种技术来分析内皮一氧化氮合成酶(eNOS)的体内磷酸化位点的结果表明，这种系统对复杂样品中数量较少的多肽进行 MS/MS 分析的功能很强。

(三) 仪器的发展

1. 涂层毛细管

涂层的毛细管常常被用来分析生物样品，如碱性多肽和蛋白质等，因为它可以减小由于这些物质与毛细管内壁发生吸附作用而引起的条带变宽和峰的拖尾现象。CE 的毛细管涂层有多种。氨丙基甲硅烷化(aminopropyl silyl, APS)的熔融石英毛细管是 CE/MS 联用中使用最多的一种涂层毛细管。因为经过这种处理的毛细管内壁带上了正电荷，这对多使用酸性挥发性缓冲液进行蛋白质和多肽分析的 CE/MS 来说非常合适。线性聚丙烯酰胺(linear polyacrylamide, LPA)和聚乙烯基乙醇(polyvinyl alcohol, PVA)涂层的毛细管被广泛地用在 CIEF 实验中。经过这种方法涂层后的毛细管表面是中性的，其中的电渗流几乎为零。另一种经过亲水性衍生的毛细管也常常用于 CE/MS 联用之中，用这种经过表面疏水性改变的毛细管进行分离常常可以获得意想不到的效率提升。另外，使用动态非共价涂层的毛细管可用于多种物质的分离，用这种方法进行涂层非常方便而且重复性较好。最近，还有许多新型的涂层方法和试剂已经研制出来，比较有代表性的有采用陆续的多层涂层方法，以及采用新型的多聚物进行毛细管的涂层。

2. 捕集离子型质谱仪

提高 CE/MS 灵敏度的另一个途径就是发展能提供更高灵敏度的 MS 仪器。现在的 MS 中，从大气压 ESI 源到 MS 检测器的离子转移效率很低。因此，使用捕集离子型质谱仪，包括 FTICR-MS 和 IT-MS，可能会对提高灵敏度有益。

IT-MS 能比柱型(beam-type)MS 提供高得多的传输效率，它的离子捕集能力使得它在进行 MS^n 分析时有独特的优势。已有报道一些商品化仪器(如 Finnigan LCQ 等)进行了 MS^{10} 的实验。最近有不少 CE/IT-MS 联用的研究报道[38]。

FTICR-MS 具有别的 MS 仪器所不能比拟的高灵敏度、高精确度(ppm 水平)、超高的分辨率($>10^6$)和进行串联质谱分析的强大潜力(FTICR-MS 的非解构性检测方法使得它可以反复地对某一捕集的离子群进行解离和分析其解离产物，从而可从有限的样品中得到更丰富的结构信息)。近年来 FTICR-MS 发展很迅速，它在与其他离子源的接口、离子捕集室中的离子操作技术以及数据处理方法等方面都取得了很大的进展。Belov 等[39]人报道 ESI-FTICR-MS 对蛋白质的检测灵敏度已经达到 zeptomole(10^{-21}mol)水平。Ramjit 等[40]人用 FTICR-MS 通过分辨单一同位素峰和 C^{13} 同位素峰成功地对一个有 19 个成分的十二肽库进行了鉴定。正是由于 FTICR-MS 的这种超高分辨率，使得他们在进行这项对别的 MS 仪器来说可能是难以想像的工作时显得游刃有余。Smith 的研究小组在发展 ESI-FTICR-MS 以及将它与 CE 联用方面有极大的贡献。Hofstadler 等[41]人首次利用

ESI 接口将 FTICR-MS 与 CE 实现了联用。Severs 等[42]人还利用该系统对 CZE、CITP 及 CIEF 与 FTICR 仪器的联用作了研究。在用 CIEF/FTICR-MS 来分析红细胞裂解产物中的蛋白质时，尽管其中血红蛋白的量比碳酸酐酶要高出 100 倍以上，这两种 p*I* 值差异只有 0.27 的蛋白质还是得到了完全而且高效的分离。Hofstadler 等[43]人甚至使用 CE/ESI-FTICR-MS 技术和显微进样的方法鉴定了单个红细胞中的血红蛋白。

3. 飞行时间质谱仪

TOF-MS 仪器的最大特点是极高的扫描速度及其高工作周期所提供的高灵敏度，它可以从单个的进样脉冲中记录下全 m/z 范围的质谱图，这对于检测一般峰宽较小的 CE 流出物质十分有利。而且，TOF 检测器还可以通过反射器式(reflectron)的设置来进一步提高分辨率。同时，这种检测器相对较简单而且更便宜，对许多实验室来说极具吸引力。

TOF 检测器尤其适合与那些脉冲式离子化的离子源配合使用，但它在与容易和 CE 实现在线联用的 ESI 离子源联合使用也同样极具潜力。为了达到 ESI-TOF-MS 分析的最佳灵敏度，最好使 ESI 的离子流以正交的方式进入 MS 仪器。Lubman 的小组在实现 CE/ESI-TOF-MS 联用方面做了许多尝试，他们发展了一种离子阱贮存/反射型(ion trap storage/reflectron, ITSR)TOF-MS 仪器。这种系统利用一个离子阱作为 TOF 分析前的前端离子贮存设备，它把连续的离子流转换成脉冲离子流。而且，可以以相对较低的 pulse-out 频率使离子阱获得高工作周期而不会损失灵敏度和质量分辨率。Wu 等[44]人利用 ITSR-TOF-MS 与 CZE 联用来鉴定蛋白质，这种 MS 仪器还可以和密封式和开放式的 CEC 方法联用来分析蛋白质和多肽混合物。最近，Jin 等[45]人使用 ITSR-TOF-MS 仪器结合 SWIFT 技术对 CE/μESI-TOF-MS/MS 作了进一步的研究，取得了极好的结果。

五、应　　用

(一) 蛋白质

CE/MS 在蛋白质研究中的重要方面就是被用来对溶液或凝胶中的蛋白质进行鉴定。利用 MS 肽质谱(peptide mass mapping)所产生的各个酶解肽段的质量数信息集合或通过 CID 获得的酶解肽段的 MS/MS 图谱，就可以分别利用不同的算法来进行蛋白质序列库的搜索，从而得到蛋白质的鉴定结果[46]。其中，CID 质谱图包含有肽段的全部或部分序列信息。因为在不同的 CID 条件下，肽段会在肽键及侧链处发生断裂，从而产生一系列的序列特异性离子(图 10.3)。

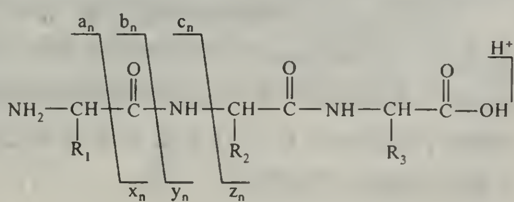


图 10.3　多肽经 CID 生成的碎片离子的命名方法

N 端带有电荷的不同碎片离子分别称作 a_n、b_n或 c_n离子；C 端带有电荷的不同碎片离子分别称作 x_n、y_n或 z_n离子。在低能量条件下主要生成 b 和 y 类型的碎片离子

CE/MS 系统用于蛋白质鉴定最大的优点就是其高灵敏度和比 LC/MS 更短的分析时间。Figeys 等人对使用 CZE/μESI-MS/MS系统进行蛋白质鉴定作了比较深入的研究。该系统在使用 20μm 内径的毛细管及无鞘流液接口进行成功的蛋白质鉴定只需

50 fmol 的酶解样品[3]。后来，他们将 SPM 技术用于 CZE/μESI-MS/MS 系统后，灵敏度进一步得到提高，进行 MS/MS 分析的浓度检测极限达到 300 amol/μl，所消耗样品量小于 10 fmol。他们用这种方法对 2-DE 凝胶内的酵母蛋白质进行了鉴定，充分显示了这种系统在大规模蛋白质鉴定方面的巨大潜力。另外，Valaskovic 等[4]人通过使用内径仅有 5μm 的毛细管进行 CZE/nESI-FTICR-MS 并首次报道检测到了 9amol 的碳酸酐酶，分辨率达到 60 000 以上。Katayama 等[47]人最近对有机溶剂辅助的 CE/nESI-MS/MS 技术用于高灵敏度蛋白质鉴定作了介绍。

对蛋白质鉴定最大胆的一种设想可以说就是同时鉴定复杂混合物中的所有蛋白质了，因为这对分离系统的分辨率和检测系统的灵敏度、精确度和动态范围要求极高，简直可以说是“Mission impossible”。但是，如此纷繁复杂的分析鉴定任务也可以通过“复杂”的 CE/MS/MS 系统得以解决。Tong 等[48]人通过一种设计巧妙的 SPM/多步洗脱-CE/MS/MS系统对含有 75 个酵母核糖体蛋白质的复杂混合物进行了分析。结果表明，高达 90％的蛋白质可以通过使用实验中所获得的 CID 图谱进行数据库搜索而得到鉴定。SPM 系统实际上还通过使用有机相的分步洗脱梯度增加了除 CE 之外的额外分离效果，所以使得 66 种蛋白质能被同时鉴定。

蛋白质分析的另一个主要方面——蛋白质修饰的分析将为揭示其结构与功能的关系提供很大的帮助。我们通过 CZE/ESI-IT-MS/MS 联用对 β-酪蛋白中的丝氨酸磷酸化进行了分析[38]。通过 SIM 模式，可以很方便地找到磷酸化肽，并且利用其 CID 图谱对磷酸化肽的结构作进一步的确认。Cao 等[49]人则利用固定化金属亲和层析与 CE/ESI-MS^n的在线联用来确定蛋白质的磷酸化位点。对糖蛋白的研究也有许多 CE/MS 的成功应用。Yeung 等[50]人发展了一种利用 CE/TQ-MS 直接对含高甘露糖型的糖蛋白进行分析的方法，不但可以分辨不同的糖蛋白糖型，还可以通过对完整蛋白质的源内裂解以产生特征性报告离子来观察糖链的存在。另外，Boss 等[51]人还通过 HPLC 和 CE/MS 对重组人红细胞生成素进行了结构分析。糖基化位点、糖链结构都可以得到确认；位点特异性的微观不均一性也可以直接估定；糖基化位点的占用率还可以进行定量测定。用这种方法可以获得 98％以上的序列覆盖率。

蛋白质的高级结构变化如蛋白质的还原和氧化态结构差异[52]、蛋白质的重折叠[53]等都可以利用不同的 CE/MS 联用形式来进行研究。

另外，CIEF/MS 联用还提供了除 2-DE 以外进行高分辨率蛋白质组研究的另一种有力手段。Lee 和 Smith 的小组在这方面作了很出色的工作[54~57]。尤其值得一提的是，Smith 小组通过其 CIEF/μESI-FTICR-MS 系统的高灵敏度和高精确度并结合选择性同位素标记的方法，还可以对蛋白质的表达水平作出评估，从而实现了真正意义上的 CIEF/MS蛋白质组研究[57, 58]。

（二）多　肽

CE/MS 对多肽的研究主要集中在对被修饰多肽和生物活性肽进行鉴定这两个方面。对比 LC/MS 而言，CE/MS 不会丢失那些极性很强的寡肽。通常对多肽的分析都是利用 CZE/MS 联用模式来进行的，但最近 Shen 等[59]人发展的高效 CIEF 也可成功地用于多肽混合物分析。另外，Barroso 等[60]人还发展了一种无鞘流液预浓缩 CZE/MS 方法用于

多肽分析。

对于磷酸化肽的分析，固定化金属亲和层析与 CE/MS 在线联用是一种很好的方法[61]，可进行数据依赖性自动电压微调的 SPE-CE/μESI-MS/MS 系统还可检测到复杂多肽混合物中的微量磷酸化肽[37]。Soudan 等[62]人将 CZE 与 MALDI-MS 结合对一个苏氨酸/丝氨酸含量丰富的 MUC5AC 十六肽的体外 O-糖基化情况进行监测，取得了很好的结果。另外，Forest 等[63]人还通过 CE/MS 确定了一种海银莲花毒素的 C 末端氨基酸为游离羧基结构。

将 CE/MS/MS 用于天然多肽的分析不但可以确定其组分的多少，还可以得到各组分的结构信息，是一种方便快捷的方法。Perkins 等[64]人用 CE/MS/MS 结合 SIM 的方法从黑树眼镜蛇(black mamba)的毒液中确定了 70 多种多肽的存在，它们的分子质量范围在 6 000～9 000 之中。Ensing 等[65]人通过 CE/MS 将神经肽 Y 与它的两个片段分离。而且，神经肽 Y 的分解产物由于只有一个氨基酸残基的差别而不能被 CE-UV 检测所区分，但 CE/MS 却可以把这些分解产物分别鉴定出来。CE 和 MALDI 的非在线联用同样在多肽分析中起着很重要的作用：Zhang 等[66]人使用这种技术对 E 肽在大脑内的代谢过程进行了监测；Ivanov 等[67]人则鉴定了多肽中的异常氨基酸残基(修饰化的和非编码的)。值得一提的是，Lynen 等[68]人还发展了一种用万古霉素结合 ACE/MS 的方法来分析寡肽混合物。

(三) 非共价复合物

CE/MS 是研究生物分子间非共价相互作用的一种极具价值的技术，因为它集分离、鉴定和定量分析于一体，而且快速准确。最近有许多利用 CE/MS 技术研究非共价复合物的报道。Lyubarsakya 等[69]人利用 ACE/MS 方法来进行内啡肽线性抗原决定簇的定位。他们首先用 CE/MS 分析内啡肽的胰蛋白酶酶解产物，然后再将这种酶解产物与抗内啡肽的单抗混合进样进行相同条件的 CE/MS 分析。这样，与单抗结合的肽段就会从它原来出现的地方消失。把这个肽段分离后用 ACE/MS 进行分析可以确认其中抗原决定簇的存在。最后通过 ACE/MS 对该肽段的一系列按顺序截短产物进行分析就可以精确地确定线性的抗原决定簇的位置。这种方法每次分析所消耗的抗体和酶解产物都在 fmol 水平，而且快速，容易实现自动化。另外，他们还用 CIEF/MS 方法进行了 *Src* SH2 结构域的高亲和性配体的筛选[70]。Tuffal 等[71]人用 CZE/MS 方法研究了分枝杆菌的甲基葡萄糖脂多糖与棕榈酰辅酶 A 形成的非共价复合物，发现它们在复合物中的化学计量为 1∶1。另外，Gysler 等[72]人还利用这种联用技术对白介素 6 的二聚体化过程进行了监测。

第二节　材　　料

一、试剂和溶剂

① 所有缓冲液均用经过 Millipore(Bedford, MA, USA)系统处理过的水配制。

② 甲酸：购自 Sigma(St. Louis, MO, USA)公司。

③ 冰醋酸：购自 Sigma(St. Louis, MO, USA)公司。

④ 氨水:购自 Sigma(St. Louis, MO, USA)公司。

⑤ 碳酸氢铵:购自 Sigma(St. Louis, MO, USA)公司。

⑥ 二硫苏糖醇(DTT):购自 Sigma(St. Louis, MO, USA)公司。

⑦ 碘代乙酰胺(IAA):购自 Sigma(St. Louis, MO, USA)公司。

⑧ β-巯基乙醇:购自 Merck(Schuchardt, Germany)公司。

⑨ 乙腈(ACN):购自 Fisher(Fair Lawn, NJ, USA)公司。

⑩ 甲醇(MeOH):购自 Fisher(Fair Lawn, NJ, USA)公司。

⑪ 三氟乙酸(TFA):购自 Merck(Schuchardt, Germany)公司。

⑫ 1,5-二甲基-1,5-二氮十一亚甲基聚溴化物(polybrene):购自 Sigma(St. Louis, MO, USA)公司。

⑬ 表面活性剂 Brij® 35 和 SM-2 吸附小珠:购自 Bio-Rad(Hercules, CA, USA)公司。

⑭ 两性电解质 pharmalyte pH3.5～10 和 pH5～8:购自 Amersham Pharmacia Biotech (Uppsala, Sweden)公司。

二、多肽和蛋白质

① Met-脑啡肽:购自 Sigma(St. Louis, MO, USA)公司。

② Leu-脑啡肽:购自 Sigma(St. Louis, MO, USA)公司。

③ 人工合成鲑鱼降钙素(s-CT):生化所崔大敷研究员惠赠。

④ ^{125}Ser 突变的重组人白细胞介素 2(S125-IL-2):上海医药工业研究院提供的基因工程产品。

⑤ 血红蛋白 A(HbA):购自 Sigma(St. Louis, MO, USA)公司。

⑥ 镰状红细胞血红蛋白(HbS):购自 Sigma(St. Louis, MO, USA)公司。

⑦ 牛 β-酪蛋白(β-casein):购自 Sigma(St. Louis, MO, USA)公司。

⑧ 牛胰 RNase B:购自 Sigma(St. Louis, MO, USA)公司。

⑨ 竹叶青蛇毒 C-型凝集素(TSL):基本参照梁雪莲等[73]人的方法获得。

⑩ 细胞色素 c:购自 Pierce(Rockford, IL, USA)公司。

⑪ 马肌红蛋白:购自 Sigma(St. Louis, MO, USA)公司。

⑫ 牛 β-乳球蛋白 A:购自 Sigma(St. Louis, MO, USA)公司。

⑬ 牛碳酸酐酶Ⅱ:购自 Sigma(St. Louis, MO, USA)公司。

三、酶

① TPCK-胰蛋白酶:本实验室自制。

② 蛋白磷酸酶 1 催化亚基(PP1):购自 Roche(Mannheim, Germany)公司。

③ Lys-C 蛋白酶:购自 Sigma(St. Louis, MO, USA)公司。

④ Glu-C 蛋白酶:购自 Pierce(Rockford, IL, USA)公司。

⑤ α-甘露糖苷酶:购自 Sigma(St. Louis, MO, USA)公司。

⑥ *N*-糖苷酶 F:购自 Roche(Mannheim, Germany)公司。

四、仪　　器

熔融石英毛细管和毛细管卡座购自 Beckman(Fullerton, CA, USA)公司。线性聚丙烯酰胺(linear polyacrylamide, LPA)涂层的毛细管为 Bio-Rad(Hercules, CA, USA)公司产品。快速脱盐柱为 Amersham Pharmacia Biotech(Uppsala, Sweden)公司产品 PD-10 预装柱。

毛细管区带电泳(CZE)和毛细管等电聚焦(CIEF)在 Beckman P/ACE Station 5500 型毛细管电泳仪上进行。熔融石英毛细管全长 80cm,入口端到 UV 检测器的距离为 20cm,内径为 50μm,外径为 360μm。质谱检测在 Finnigan MAT(San Jose, CA, USA)公司的离子阱质谱仪 LCQ 上进行,数据采集使用 LCQ Navigator 1.2 软件包,数据解析使用 LCQ BioExplore 1.1 软件包。

第三节　方　　法

一、胰蛋白酶酶解

100μg 蛋白质溶于 30μl NH_4HCO_3(100mmol/L, pH7.8)缓冲液中。先用 DTT(终浓度为5mmol/L)在 50℃ 条件下还原 15min,冷却至室温后,加入 IAA(终浓度为 10mmol/L)在室温下再反应 15min(避光)。用 NH_4HCO_3缓冲液稀释到 100μl,加入浓度为 1g/L 的胰蛋白酶溶液 2.5μl,37℃ 保温 4h。补加酶 2.5μl 后,37℃ 继续保温 18h。偶尔取出轻摇。酶解完成后取出置于 −20℃ 保存。

二、血红蛋白的脱盐和纯化

由于商品化的血红蛋白含有碳酸氢铵及 50%左右的非离子稳定剂,所以在进行蛋白质酶解之前,首先需要进行纯化。快速脱盐柱 PD-10 先用 0.6% HAc 溶液平衡至少 10 倍柱体积(25ml)。HbA 和 HbS 均溶于 2.5ml 水中,浓度约为 10 g/L。待柱内液面正好没入柱中填料时,将 2.5ml 样品全部加入。待液面再次没入柱中填料时,用 0.6% HAc 溶液进行连续的洗脱,收集前 3.5ml 洗脱物。冷冻干燥后将样品置于 −20℃ 保存。

三、蛋白磷酸酶去磷酸化

β-酪蛋白 1g 溶解于 1ml 去磷酸化溶液[25mmol/L 三乙醇胺(TEA), 200mmol/L NaCl, 0.1mmol/L EGTA, 0.5mmol/L $MnCl_2$, 0.1% β-巯基乙醇(m/V), 0.03% Brij® 35(m/V), 用 HCl 调节 pH 至 7.5]中。加入 5mU/μl 的 PP1(溶于 50% 甘油(V/V), 25mmol/L TEA-HCl, 200mmol/L NaCl, 0.1mmol/L EGTA, 0.1% β-巯基乙醇(m/V), 0.03% Brij® 35(m/V), pH 7.5)溶液 5μl,37℃ 保温 4h。反应完成后冷至室温。

20mg SM-2 吸附小珠先用 50ml 甲醇洗 10min,重复 2 次。加入 50ml 水洗 10min,重复 3 次。把水吸去,加入 1ml 已反应完的溶液,室温下摇动 2h。吸取上清,加入另一份已

经用同样方法处理过的 SM-2 吸附小珠，室温处理 2h。

将反应液用 0.6% HAc 溶液稀释到 2.5ml，用 PD-10 预装柱脱盐。收集的 3.5ml 液体冷冻干燥后，置于 -20℃ 保存。

四、蛋白质二硫键的羧酰胺甲基化及产物的 HPLC 纯化和蛋白酶酶解

0.2mg RNaseB 溶于 80μl 30mmol/L NH_4HCO_3 缓冲液（pH7.8）中，加入 50mmol/L DTT 10μl，50℃ 反应 15min，再加入 100mmol/L IAA 10μl 在 37℃ 反应 15min。

0.2mg TSL 溶于 80μl 30mmol/L NH_4HCO_3 缓冲液（pH7.8）中，加入 0.5mol/L DTT 10μl，50℃ 反应 1h，再加入 1mol/L IAA 10μl 在 37℃ 反应 30min。

HPLC 为 HP-1100 型，使用 C8 反相柱（RP-300，5μm，30mm×2.1mm I.D.，ABI 公司产品）；A 液：0.01% TFA-H_2O；B 液：0.01% TFA-ACN；流速：200μl/min；洗脱梯度：0min～2min，0% A 液，2min～15min，0～100% B 液。

100μg 经纯化的羧酰胺甲基化蛋白质直接溶于 100μl 30mmol/L 的 NH_4HCO_3 缓冲液（pH7.8，胰蛋白酶或 Lys-C 蛋白酶酶解）或 100μl 30mmol/L 的 NH_4Ac 缓冲液（pH4.0，Glu-C 蛋白酶酶解）中。加入浓度为 1 g/L 的蛋白酶溶液 2.5μl，37℃ 保温 4h。补加酶 2.5μl 后，37℃ 继续保温 18h。偶尔取出轻摇。反应完成后置于 -20℃ 保存。

五、*N*-糖苷酶 F 酶解

90μl 蛋白质的蛋白酶酶解产物中加入 1 U/μl 的 *N*-糖苷酶 F 水溶液 10μl，37℃ 保温 4h。反应完成后置于 -20℃ 保存。

六、α-甘露糖苷酶酶解

90μl 蛋白质的蛋白酶酶解产物中加入 0.1 U/μl 的 α-甘露糖苷酶溶液（3.0mol/L $(NH_4)_2SO_4$，0.1mol/L $ZnSO_4$，pH7.5）10μl，37℃ 保温 6h。反应完成后置于 -20℃ 保存。

七、正离子非共价涂层毛细管电泳

正离子非共价涂层方法基本参照 Kelly[74] 和 Wiktorowicz 等[75] 人的方法进行，简要概述如下：毛细管用 1mol/L NaOH 溶液冲洗 15min，再依次用水、0.1mol/L HCl 溶液和水各冲洗 5min。然后用涂层液[含有 5%（*m*/*V*）polybrene 和 2%（*V*/*V*）乙二醇的水溶液]冲 5min。最后用 2mol/L 甲酸溶液冲 5min。电泳缓冲液为 2mol/L 甲酸溶液，分离电压为 -30kV，在 200nm 波长下检测。在每次实验结束后需要用同样的方法重新涂层。

八、毛细管电泳-质谱联用

新鲜毛细管内壁的硅醇基团通过以下方法来激活：1mol/L 的 HCl 冲洗 15min，水冲

洗 15min，再用 1mol/L 的 NaOH 继续冲洗 15min，最后用水冲 15min。在不同的电泳过程之间，毛细管用 0.1mol/L 的 NaOH 再生 5min，再用水冲洗 5min，最后用电泳缓冲液平衡 5min。

CE/MS 接口采用鞘流液接口。毛细管出口端不导电的高聚物涂层被烧去 2～3mm，裸露出石英毛细管，以利于鞘流液接口处的电接触。毛细管的入口端置于缓冲液储瓶中。为了将由于重力作用引起的液体流动带来的影响降到最低，整个毛细管电泳仪器被放在一辆可以精细调节高度的小车上，并仔细调节仪器的高度以使毛细管入口端缓冲液储瓶与毛细管出口端所在的质谱仪离子源喷雾口处于同一水平上。鞘流液（60%乙醇，1%甲酸，39%水）用 500μl 的 Unimetrics 注射针筒，由质谱仪上的注射泵推动通过离子源上套在毛细管外部的一根不锈钢细管引入，流速为 2.00μl/min。由于 CE/MS 中毛细管的流量很低，一般＜1μl/min，所以不使用流量较大的 LC/MS 中所必需的鞘流气（sheath gas）和辅助气（auxiliary gas）。这样，毛细管末端突出在不锈钢细管外部分的长短对喷雾的效果好坏就极为关键。在 Finnigan 公司的 CE/MS 接口上，可以通过一个千分尺移动头来作毛细管末端位置的精细调节，一般突出长度约为 1mm。CE/MS 接口处的不锈钢细管上加上 4.25kV（对比质谱仪毛细管入口处）的正向喷雾电压。毛细管电泳分离通过在毛细管进样口端加上正向高压来进行。

质谱仪在联机状态下先用血管紧张肽做校正，以使仪器对毛细管电泳的出口情况做出相应的参数调整以获得最佳的检测灵敏度和分辨率；然后再用 Finnigan 的 Ultramark 做 Calibration 来调节质谱仪的精确度及在不同扫描方式下的参数以得到最佳的操作性能。其他主要操作参数为：质谱仪入口处加热毛细管的温度设为 180℃，其加速电压设为 13V；Tube lens offset 为 55V；正离子模式检测。在 MS/MS 实验中设置 CID 碰撞能量为 45%，所要求的最小信号计数为 10^5。所有实验均采用自动控制（auto gain control，AGC）模式进行，即在质谱仪的扫描设置中包含三个阶段：首先是质谱仪在 m/z 400～2 000 的范围内作全扫描（full scan），检测所有经过加热毛细管进入质谱仪的离子；随后再对信号计数超过 10^5（可另行设置）的丰度最大的离子，在 m/z 很小的范围内进行精度很高的放大扫描（zoom scan），可用于确定离子的带电荷情况；最后质谱仪对已做过放大扫描的离子在离子阱中做 MS/MS，并将生成的所有碎片离子记录下来。随后质谱仪又回到全扫描状态，依次循环下去。若质谱信号未超过所设定的最小信号计数，则质谱仪将一直保持在全扫描状态。这三个阶段可在 1s 之内完成，从而能够保证获得足够的质谱信息。

九、蛋白质序列数据库搜索

蛋白质序列数据库的搜寻通过 SEQUEST 软件进行。对于经过 MS/MS 分析的质谱图，SEQUEST 将给出一个表格，其中包含从数据库中抽出的最可能的肽段，以及它们的 Xcorr 相关系数和 DelCn 相关因子。Xcorr 相关系数表示，将实验中获得的 MS/MS 质谱图与数据库中相应肽序列的预测的碰撞诱导解离质谱图所作比较（match）的质量；DelCn 相关因子表示，最匹配的肽段与其次的肽段之间的差异。一般 Xcorr 相关系数在 2 以上即表明，从数据库中检索得到的肽段与实验中获得的质谱图得到了很好的匹配；DelCn 相关因子在 0.1 以上则表明，搜寻所得的最佳肽段与其次的肽段有显著的区别。

对一个蛋白质酶解产物进行分析所得到的 MS/MS 质谱图,可以不加以任何分析就直接对整个质谱图用 SEQUEST 软件进行检索;同样,也可以先经过一定的分析后,再选取感兴趣的肽段的 MS/MS 质谱图进行单独检索,这样可以大大缩减数据库检索所需的时间。对于检索出的来自同一个蛋白质的肽段,SEQUEST 将把它们归为一类,并把所检索出的蛋白质按其总加权值由高到低对它们进行排序。在此,SEQUEST 采取了一个简单的算法来作为判断的指标,即在所有从数据库中检索出来的肽段中,被认为是与该蛋白质的理论酶解肽段匹配最佳的肽段分值为 10,处于第二位的分值为 8,依次类推,第三位的为 6,第四位的为 4,第五位的为 2。最后通过各个蛋白质所获得的总加权值的高低,来判断所分析的蛋白质最可能是哪一种。

数据库搜索的参数设置为:蛋白质酶切特异性为胰蛋白酶或 Lys-C 蛋白酶,或 Glu-C 蛋白酶;选择单同位素 b 离子和 y 离子;母离子质量范围为 300～4 000;差异性修饰设为 Cys 羧酰胺甲基化(质量数增加 57.0),Met 氧化(质量数增加 16.0),静态修饰设为无;搜索所用的数据库为 SwissProt3.9 版及从中单独抽取出来的部分数据库。

十、特殊肽段的选择性离子监测

为了进一步确定蛋白质中的点突变情况,需要找到带有突变位点的特殊肽段。通过选择性离子监测(extracted ion monitoring, EIM)方法可以对分子质量已知的目标肽段进行整个质谱图范围内的搜索以查找该肽段的出峰时间,然后即可根据所找到肽段的 CID 质谱图进一步确定其中突变位点的情况。在 BioExplore 数据分析软件中输入目标肽段的可能质荷比数值即可实现选择性离子监测。

十一、毛细管等电聚焦-质谱联用

内壁经 LPA 涂层的毛细管长 80cm,内径 50μm,外径 360μm。聚焦阳极电解质为 1%乙酸溶液,聚焦阴极电解质为 1%氨水溶液;移动阳极电解质为 0.5%乙酸溶液,移动阴极电解质为含有有机相的乙酸溶液(0.25/75/25, HAc/MeOH/H_2O, $V/V/V$)。蛋白质直接溶解于含有 1%(m/V)两性电解质的水溶液中,浓度约为 0.1mg/ml。

毛细管电泳仪置于一个可以上下移动的平台上。在等电聚焦开始前,先仔细将毛细管入口端的位置调节至与 ESI 喷雾口处于同一水平位置。用聚焦阴极电解质作为鞘流液,流速始终保持在 2μl/min,同时打开鞘流气,流速为 20 个相对单位,以防止鞘流液在针管口的聚集。在将整根毛细管充满样品和两性电解质的混合物后,小心将毛细管的出口端缩入到鞘流液针管的内侧约 1.5mm 处,这样就在鞘流液针管的末端形成了一个微型的聚焦阴极电解质溶液的贮池。继续让鞘流液冲洗约 5min 以去除可能残余在毛细管末端的样品和两性电解质,然后将鞘流气流速减小到 7 个相对单位,使得鞘流液可以恰好在针管的末端稳定地形成球面型液滴。毛细管的入口端浸没于聚焦阳极电解质中,而出口端则位于以鞘流液方式引入的聚焦阴极电解质中。这时加上 30kV 的电压开始聚集(毛细管中的电位差为 375V/cm)。聚焦约 40min 后,电流不再下降,表示聚焦已经完成。此时

关闭聚集电源,迅速将毛细管位置重新调整到突出于鞘流液针管外部约 1mm 处,再迅速将鞘流液换成移动阴极电解质,入口端电解质换为移动阳极电解质。再迅速将整个毛细管电泳仪的位置向上移动 8cm(这是为了将重力引起的移动与由于移动阳极电解质替换引起的条带移动结合起来,以加快聚焦条带移动的速度),在毛细管的两端加上 + 30kV 的电压,ESI 离子源的喷雾电压设为 + 4.25kV(此时毛细管中的电位差为 322V/cm),随即开始对样品条带移动情况的 MS 在线检测过程。这种实现 CIEF/ESI-MS 联用的方法称为“活动式毛细管位置调整法”。

质谱检测条件为:加热毛细管的温度为 180℃,加速电压为 24V;Tube lens offset 为 - 5V;光电倍增管在 - 950V 条件下检测;Full scan 范围为 m/z 1 100～2 000;以正离子模式检测。

第四节 结　果

一、多肽和蛋白质的分析与鉴定

首先用 CZE/MS/MS 对 Met - 脑啡肽和 Leu - 脑啡肽的混合物进行了分析,它们的浓度分别为 133.3μmol/L 和 66.7μmol/L。缓冲液为 1%醋酸,pH3.4;在 138kPa 压力下进样 3s;电压为 30kV。图 10.4 表示了质谱仪检出的基峰图(代表在该峰中丰度最高的离子)。这两种带单个正电荷的多肽离子在经过 MS/MS 后得到了较好的碎片离子,能明确地解析出每种多肽的氨基酸序列。可以清楚地看到这两种五肽在一级结构上的区别仅仅在于 C 端的一个氨基酸残基(图 10.5)。

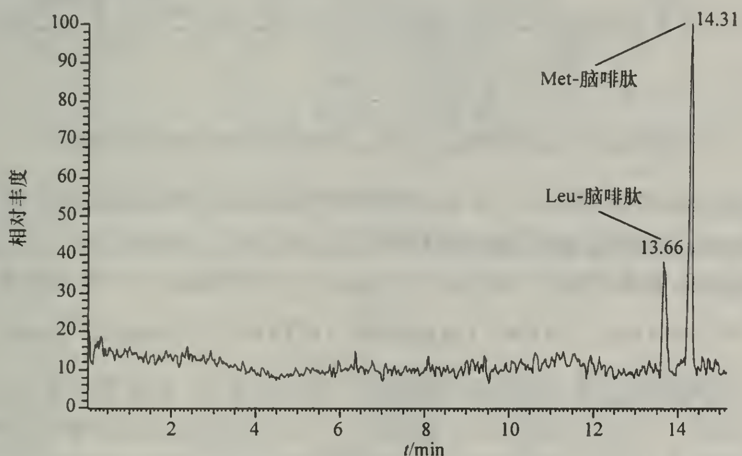


图 10.4　用 CZE/MS 方法分析 Met - 脑啡肽(1.4pmol)和 Leu - 脑啡肽(0.7pmol)的混合物

为了验证 CZE/MS/MS 系统是否适合于通过检测蛋白酶酶解产生的肽段来鉴定多肽,我们对 s-CT 的胰蛋白酶酶解产物作了分析。同样使用 1%醋酸(pH3.4)作为电泳缓冲液,s-CT 的酶解产物未经任何处理直接在 138kPa 压力下进样 5s。图 10.6 是对 s-CT 酶解产物的基峰图,每个肽段的出峰时间很短,均低于 0.4min。所有的理论酶解肽段都

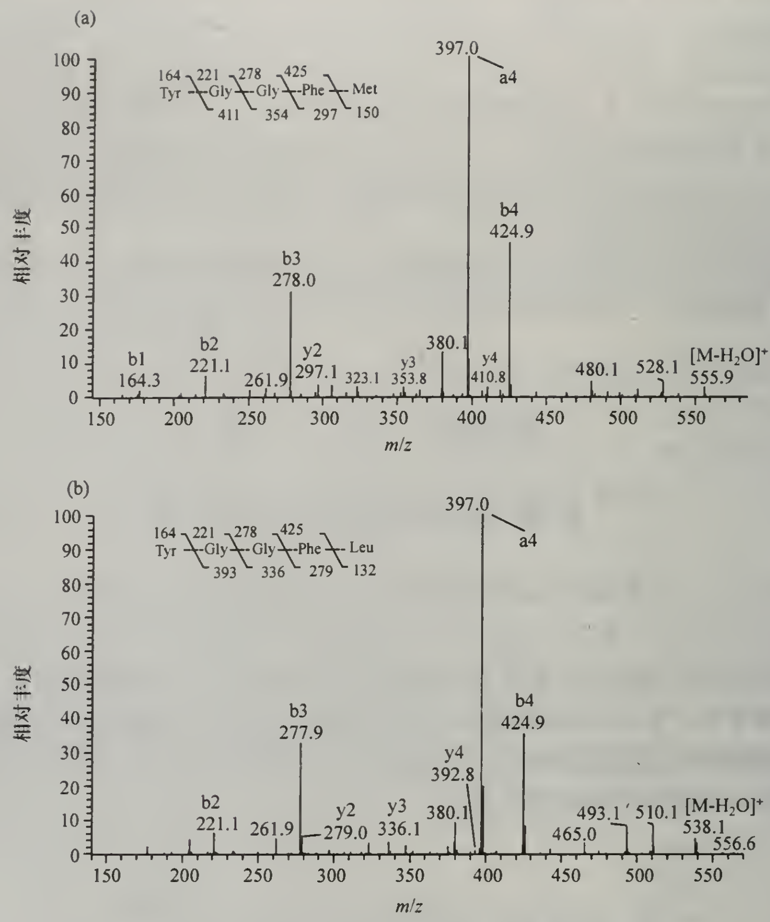


图 10.5　Met-脑啡肽(a)和 Leu-脑啡肽(b)的 MS/MS 质谱图

可由 CZE/MS 分离并得到鉴定。自动化数据库检索结果表明该多肽最可能来源于 Calcitonin 1 Precursor。其中已经被鉴定的肽段有 T1、T2 和 T3,T4 由于在人工合成的产品中 C 端为酰胺化而未能得到鉴定,但仍然可以从其 CID 图谱验证它的氨基酸序列。

^{1}C* SNLSTC* VLGK↓ LSQELHK↓ LQTYPR↓ TNTGSGT^{32}P-NH$_2$

图 10.7 是对细胞色素 c 胰蛋白酶酶解产物的基峰图,毛细管电泳在 1% 乙酸(pH 2.3)缓冲液中进行,实际分离电压为 10.75kV。几乎所有的理论酶解肽段都可由 CZE/MS 分离并检测(表 10.1)。另外,尽管一些肽段(如 T7 和 T18,T5 和 T14,以及 T9 和 T10)并未完全分开,扫描速度极快的 CZE/MS/MS 还是能够分别对它们进行鉴定。仔细观察发现,尽管肽段 T5 和 T6、T8 和 T9、T10 和 T11、T12 和 T13,以及 T15 和 T16 的氨基酸序列都只有一个 Lys 的区别,但它们的电泳迁移率却发生了很大的变化,在 CZE/MS 条件下被很好地分辨开来并分别得到了鉴定。最长的一段肽段 T11 得到了很好的 CID 图谱(图 10.8)。

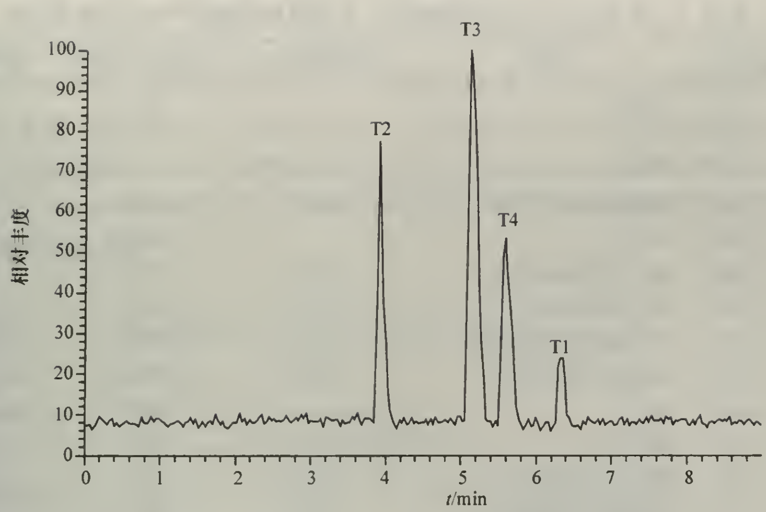


图 10.6　用 CZE/MS 方法分析 5.2pmol 的 s-CT 胰蛋白酶酶解产物(1% HAc, pH3.4)

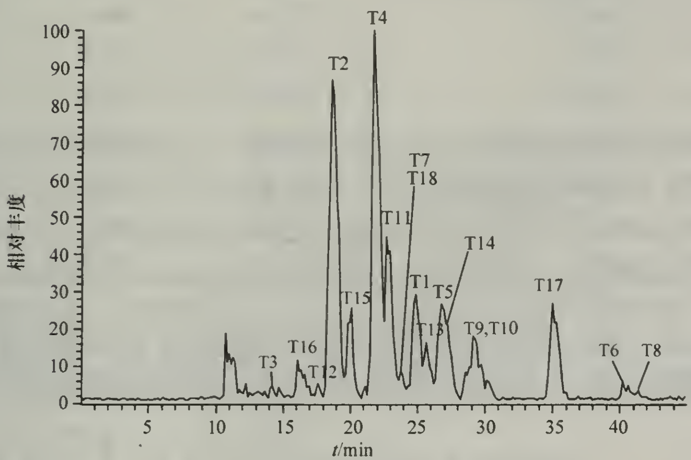


图 10.7　用 CZE/MS 方法分析 5.8pmol 的细胞色素 c 的胰蛋白酶酶解产物

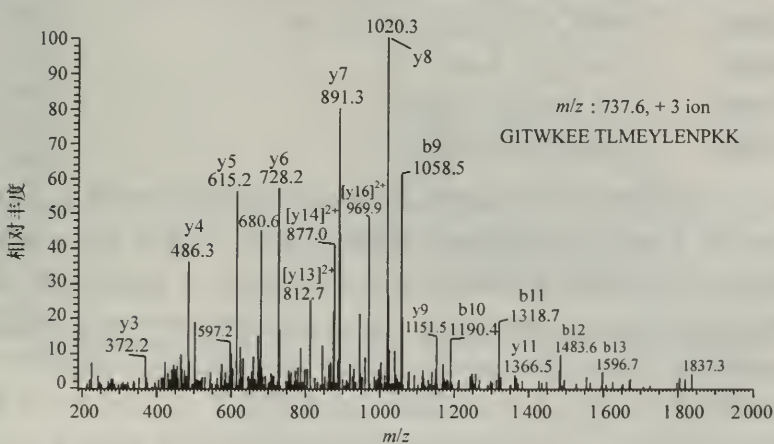


图 10.8　肽段 T11 的 MS/MS 质谱图

表 10.1 用 CE/MS 方法对细胞色素 c 的胰蛋白酶酶解产物进行肽谱分析

肽段	位置	序列	保留时间/min	理论分子质量/Da	观测分子质量/Da
	1~5	GDVEK		546.6	
	6~8	GKK		331.4	
T1	9~13	IFVQK	24.85	633.8	633.4
T2	9~22	IFVQKCAQCHTVEK	18.55	1 633.9	1 633.2
	23~25	GGK		260.3	
T3	26~38	HKTGPNLHGLFGR	14.23	1 433.6	1 433.5
T4	28~38	TGPNLHGLFGR	21.71	1 168.3	1 168.2
T5	39~53	KTGQAPGFTYTDANK	26.32	1 598.7	1 598.7
T6	40~53	TGQAPGFTYTDANK	40.17	1 470.6	1 470.2
	54~55	NK		260.3	
T7	56~60	GITWK	23.33	603.7	603.4
T8	61~72	EETLMEYLENPK	41.38	1 495.7	1 495.0
T9	61~73	EETLMEYLENPKK	28.52	1 623.8	1 623.3
T10	56~72	GITWKEETLMEYLENPK	29.10	2 081.4	2 080.9
T11	56~73	GITWKEETLMEYLENPKK	22.65	2 209.5	2 209.3
T12	73~79	KYIPGTK	17.56	806.0	805.6
T13	74~79	YIPGTK	25.64	677.8	677.3
T14	80~86	MIFAGIK	26.72	779.0	778.4
T15	88~99	KTEREDLIAYLK	20.11	1 478.7	1 478.4
T16	88~100	KTEREDLIAYLKK	16.06	1 606.9	1 607.1
T17	92~99	EDLIAYLK	35.03	964.1	963.5
T18	100~104	KATNE	23.75	561.6	561.4

CZE/MS/MS 一共生成了 29 个不同的 CID 质谱图。用 SEQUEST 软件对事先未加以任何解析的酶解产物 MS/MS 质谱图,先进行马蛋白质序列数据库的检索。结果如下:

A

```
1.CYC_HORSE     180(18,0,0,0,0){ 4 6 8 9 10 11 12 13 14 15 16 17 18 19 23
                                                          25 26 27 }
2.CPT7_HORSE    60(4,1,2,0,0){ 3 21 24 28 }
3.TRFL_HORSE    50(0,3,3,0,4){    }
4.ETBR_HORSE    46(1,2,1,3,1){ 29 }
5.GELS_HORSE    44(0,2,4,1,0){    }
```

共有 18 个 CID 质谱图与马细胞色素 c 的某些序列相匹配,它们的 Xcorr 相关系数和 DelCn 相关因子都比较高,显示出很高质量的检索结果。尤其是 DelCn 相关因子大于 0.4,这表示通过检索所得到的肽段具有比较明显的特殊性,与马细胞色素 c 的胰蛋白酶酶解产物中的某些肽段对应关系极为显著。而且,检索得到的序列较长的肽段均具有较高的 Xcorr 相关系数(如最长的两段肽段 T10 和 T11 的该项值分别达到了 5.7017 和 5.1797),说明数据库中对这些肽段所预测的 CID 质谱图与 CZE/MS/MS 实验所获得的 CID 质谱图能很好地吻合。点击相应肽段的因特网链接还可以观看该肽段的理论 CID

图谱,以供与手工解析结果相对照。马细胞色素 c 的总加权值最高,为 180,有 18 个 CID 质谱图与之显著相关(列为最佳匹配肽段),而其他几个蛋白质的总加权值都不高。仔细观察检索结果发现,蛋白质 CPT7 HORSE 的 4 个最佳肽段均来自于在不同的时间分别出现 4 次的一个质量数约为 452.2 的离子,而这极有可能就是来自系统内部的杂质离子。这说明利用 CZE/MS/MS 所产生的 CID 质谱图检索蛋白质序列数据库来鉴定蛋白质,可以得到快捷而可靠的结果[76]。

B

```
 1.CYC _ EQUAS      50(5,0,0,0,0){ 10 12 18 19 27 }
 2.CYC _ HELAS      30(0,3,1,0,0){ }
 3.CYC2 _ MOUSE     20(2,0,0,0,0){ 9 17 }
 4.CYC _ HORSE      20(2,0,0,0,0){ 16 26 }
 5.CIN6 _ HUMAN     20(0,1,2,0,0){ }
 6.CYC _ APIME      20(0,0,2,2,0){ }
 7.CYC _ ASTRU      20(0,1,2,0,0){ }
 8.AFAC _ ECOLI     20(1,0,1,1,0){ 29 }
 9.CYC _ ANAPL      20(2,0,0,0,0){ 4 11 }
10.CYC _ APTPA      20(2,0,0,0,0){ 4 11 }
```

如果所鉴定蛋白质的来源未知,则可将 CZE/MS/MS 产生的 CID 质谱图先在复杂非冗余库(如 SwissProt、OWL)中进行检索,这样一般可确定蛋白质的种类,然后再结合该蛋白质的分子质量进行比较精细的检索,就可确定蛋白质的种属。我们将马细胞色素 c 的酶解产物的 CID 质谱图在 SwissProt 3.9 版数据库进行检索后,结果排在前 10 位的蛋白质中有 8 种是细胞色素 c,只是来源不一样(见 B)。这主要是因为细胞色素 c 是一种很保守的蛋白质,来自不同种属的细胞色素 c 在经过胰蛋白酶酶解后可能会存在许多序列相同的肽段。进一步结合所测得的分子质量 12 359.0Da(图 10.9),即可判断该蛋白质是马细胞色素 c(理论分子质量为 12 360.0Da)。

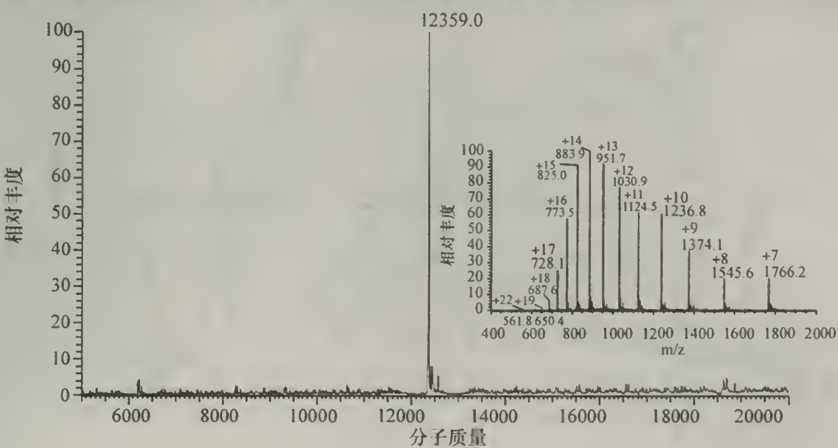


图 10.9 马细胞色素 c 的分子质量(插图表示它的电荷分布情况)

二、蛋白质点突变的鉴定

S125-IL-2 酶解产物在 138kPa 压力下进样 20s。净分离电压为 10.75kV；使用的电泳缓冲液为 1% HAc 溶液(pH2.3)。图 10.10 是 S125-IL-2 酶解产物的 base peak，一共找到了 10 个胰蛋白酶的酶切肽段，整个蛋白质序列的回收率达到了 97.8%。由于 S125-IL-2 是基因工程重组产品，所以其 N 端含有一个 Met 残基，而在天然 IL-2 中是没有这个 N 端 Met 残基的。所以在肽段 T1(MAPTSSSTK)中，将 Met 残基定位为 -1，以与天然的 IL-2 蛋白质序列相对应。

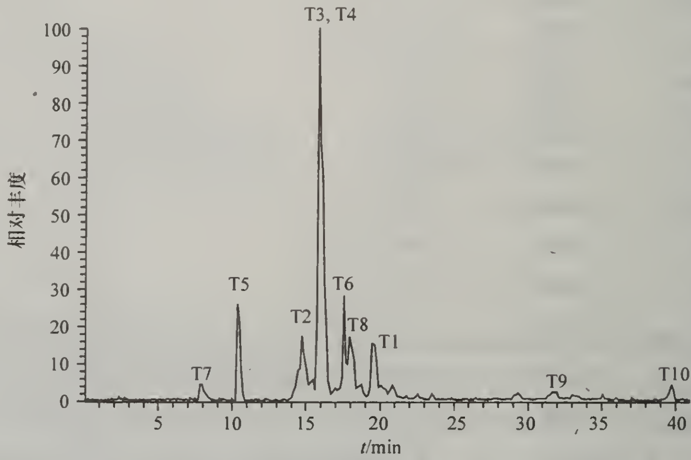


图 10.10　用 CZE/MS/MS 方法分析 4.6pmol 的 S125-IL-2 的胰蛋白酶酶解产物

为了找到带有突变位点的特殊肽段，可以利用 EIM 方法快速地进行目标肽段的搜索。T10(WITFS^{125*}QSIISTLT)的理论分子质量为 1 496.7Da，所以可以选择它的 +1 价离子(m/z = 1 497.7)进行 EIM 搜索，结果如图 10.11 所示。可见该肽段在 39.7min 时

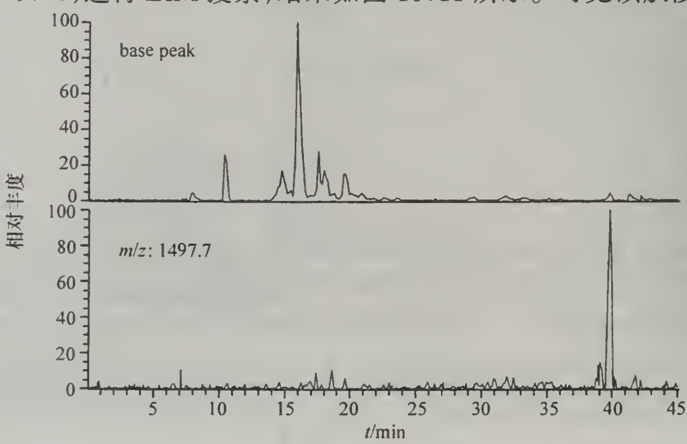


图 10.11　对质荷比为 1 497.7 的离子的选择性离子监测

出峰。然后再在该时段的质谱图中找到它的 CID 质谱图，即可实现对突变位点的鉴定(图 10.12)。从 T10 的 CID 质谱图中可以看到一个有趣的现象：T10 在经过 MS/MS 过程后生成了一系列 b 离子的连续脱水产物及肽段本身的分子连续脱水产物。这可能是因为 T10 肽段中含有较多容易脱水的 Ser 和 Thr 残基的缘故。

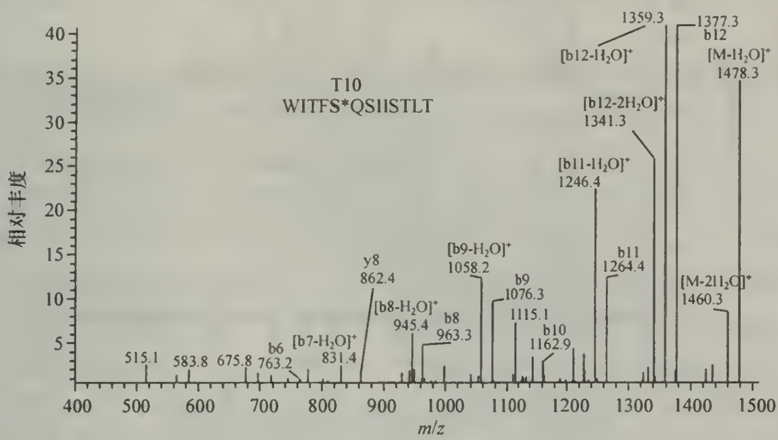


图 10.12　肽段 T10(WITFS[125*] QSIISTLT)的 MS/MS 质谱图

用同样的方法可以对 HbS 的突变位点进行鉴定。100μl HbS 的胰蛋白酶酶解产物经冷冻干燥后，用 20μl 1% HAc(pH2.6)溶解，60 000 r/min 离心 10min 后，在 138kPa 压力下进样 20s，进样量约为 11.5pmol。毛细管电泳在 1% HAc 缓冲液中进行，净分离电压为 10.75kV。由于 Hb 是一种寡聚蛋白质，分子中含有 2 个 α 链和 2 个 β 链，所以它的酶解产物中肽段很多，难以实现对每个肽段的完全分离。尽管如此，由于质谱检测的特异性、高速度和高灵敏度，利用 CZE/MS/MS 方法还是可以单独对绝大部分的肽段进行检测和鉴定。从 HbS 的基峰(图 10.13)可以看到，绝大部分肽段被找到，而且出峰时间的重现性

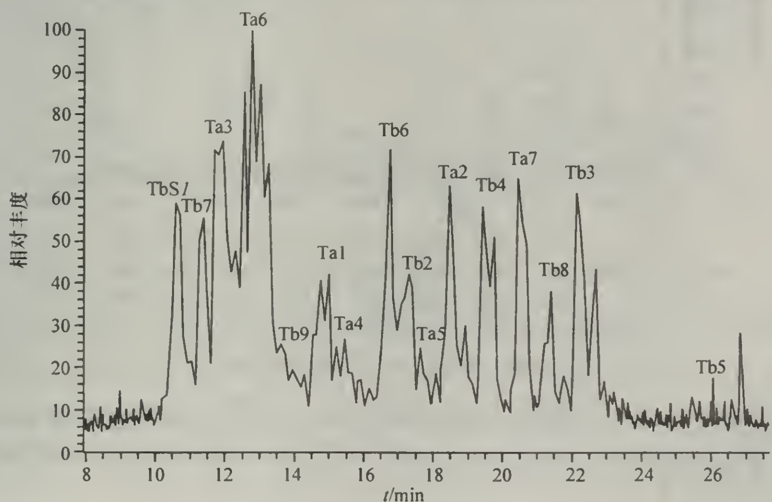


图 10.13　用 CZE/MS/MS 方法分析 11.5pmol 的 HbS

较好。对质谱图的手工解析找到了 Hb 的 α 链的 9 个理论酶解肽段中的 7 个，以及 β 链的 11 个理论酶解肽段中的 9 个，具有良好的氨基酸序列覆盖率。

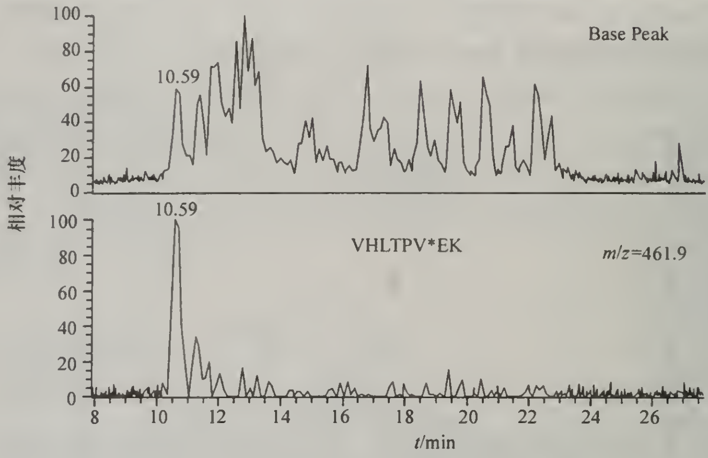


图 10.14　对肽段 TbS1 的选择性离子监测

为了进一步鉴定肽段 Tb1 中的突变位点，可以利用 EIM 方法来快速地找到它的出峰位置。TbS1 的理论分子质量为 923.1Da，所以相应地可以对其 461.9^{2+} 离子进行监测（图 10.14）。从 TbS1 的 CID 质谱图（图 10.15）可以看到，TbS1 确实发生了 β^6Glu→β^6Val 点突变。

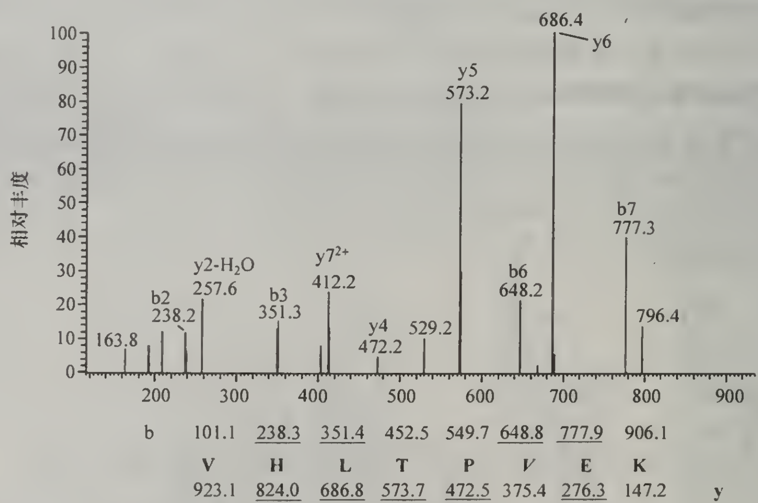


图 10.15　肽段 TbS1 的 MS/MS 质谱图

值得一提的是，发生了点突变的蛋白质同样可以用自动化的数据库检索来进行鉴定。不过含有突变位点的肽段将不能被正确鉴定，但这对整个蛋白质的鉴定影响甚微。

三、蛋白质磷酸化的鉴定

β-酪蛋白及去磷酸化 β-酪蛋白的酶解样品均在 138kPa 压力下进样 20s，相应的进样量约为 12.0pmol。净分离电压为 10.75kV；使用的电泳缓冲液为 1% HAc 溶液(pH2.6)。

β-酪蛋白的胰蛋白酶酶解产物中应该含有两段存在磷酸化位点的肽段(R^{1}ELEELNVPGEIVES^{P}LS^{P}S^{P}S^{P}EESITR^{25}和 F^{33}QS^{P}EEQQQTEDELQDK^{48})，但是由于 R^{1}ELEELNVPGEIVES^{P}LS^{P}S^{P}S^{P}EESITR^{25}含有 4 个磷酸化位点，整个肽段由于这 4 个额外的负电荷而难以在酸性条件下离子化，因而很难在正离子检测模式下被质谱仪检测到，这与文献报道是一致的[77]。利用 EIM 方法可以快速地进行目标肽段的搜索。在 β-酪蛋白中，由于含有一个天然磷酸化位点的肽段 T2(F^{33}QS^{P}EEQQQTEDELQDK^{48})的理论分子质量为 2 061.2，所以可以选择它的 + 2 价离子(m/z = 1 031.6)进行 EIM 搜索，结果如图 10.16 所示，可见该肽段在 39.6min 时出峰。根据该时段的 MS/MS 图谱，即可实现对该肽段的鉴定(图 10.17)。其中最高的信号对应于该磷酸化肽段分子的去磷酸化双电荷离子；b 离子找到 4 个，而 y 离子则找到了比较完整的一个系列(y3～y15)。这是因为，在本实验设置的 MS/MS 碰撞能量下，尽管肽键之间的断裂情况比较好，但造成磷酸化修饰的共价键却也容易发生断裂；而且磷酸化位点 Ser^{35}处于整个肽段分子的 N 端一侧，所以 y 离子较多而 b 离子较少。

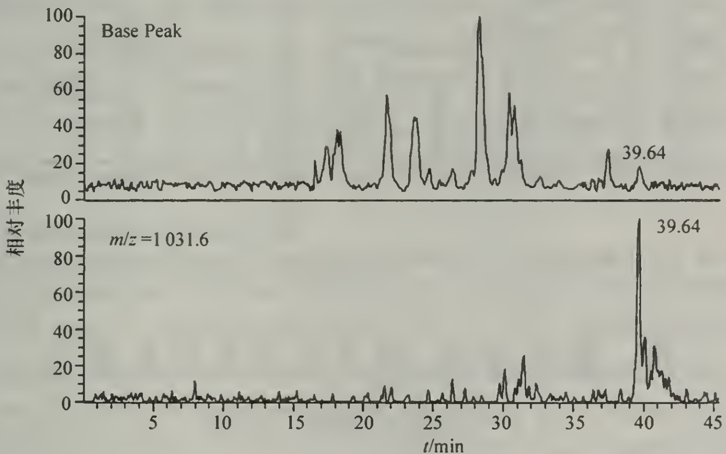


图 10.16　肽段 T2(F^{33}QS^{P}EEQQQTEDELQDK^{48})的选择性离子监测

同样用 EIM 方法可以容易地找到去磷酸化 β-酪蛋白酶解产物中的对应肽段 T2′(F^{33}QSEEQQQTEDELQDK^{48}，理论质量数为 1 982.0，搜索其双电荷离子 m/z 为 992.0)。图 10.18 是 T2′的 MS/MS 质谱图，可见其与 T2 的 CID 图谱有很明显的差别。另一段含有 4 个天然磷酸化位点的肽段的去磷酸化产物 R^{1}ELEELNVPGEIVESLSSSEESITR^{25}仍然未能得到鉴定。这可能是因为这几个磷酸化位点彼此相距太近，影响了磷酸酯酶作用的效果，导致不能或不完全去磷酸化，所以该肽段还是难以在正离子条件下得到检测。

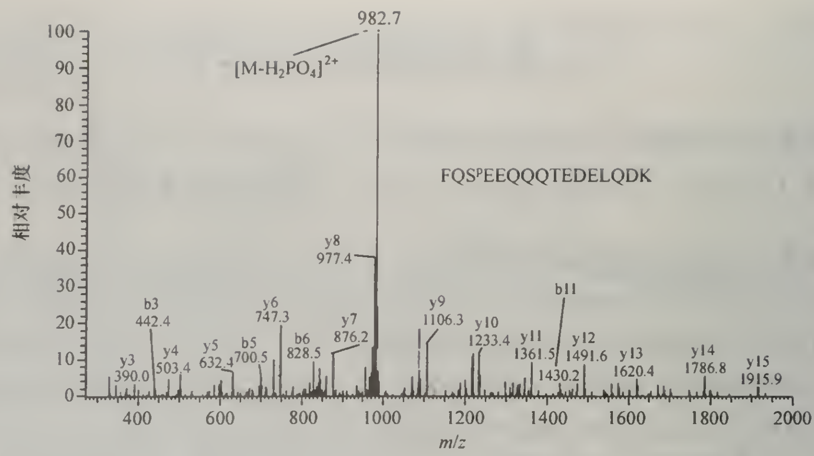


图 10.17 肽段 T2(F^{33}QS^{P}EEQQQTEDELQDK^{48})的 MS/MS 质谱图

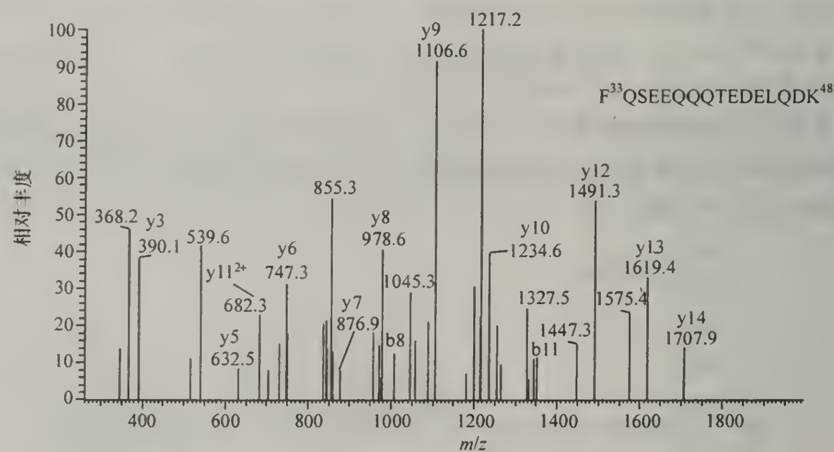


图 10.18 去磷酸化肽段 T2′(F^{33}QSEEQQQTEDELQDK^{48})的 MS/MS 质谱图

四、蛋白质 *N*-糖基化和氧化的鉴定

(一)蛋白质糖型的分辨

通常情况下,HPLC 等分离方法对于蛋白质糖型的分辨是无能为力的;但用毛细管电泳方法,选择适当的分离缓冲液和(或)涂层方法却可以分辨许多糖蛋白的糖型。在未涂层的毛细管内壁带负电荷,其中的电渗流朝向阴极;而经过正离子高聚物涂层的毛细管内壁是带正电荷的,其中的电渗流方向是朝向阳极的(电荷反转,charge reversal)。所以,在普通的极性安排下(毛细管进样端为阳极,检测端为阴极),未涂层毛细管中的电渗流与电泳迁移力方向相同,造成样品条带的扩散;在涂层毛细管中,电渗流的方向与电泳迁移力的方向是相反的,样品条带的移动速度减慢,因而分辨率得到提高。但此时样品条带的移动要靠电渗流的推动,所以,要将毛细管的进样端设为阴极,检测端设为阳极,这样样品条

带才能通过检测器得到检测，因为此时电渗流的方向是朝向阳极的。另外，通过使用强酸性的缓冲液可以将电渗流降低，从而使样品条带的移动进一步减慢，分辨率得到进一步提高。在2mol/L甲酸溶中利用正离子非共价涂层方法对RNase B的糖型分析情况如图10.19所示。

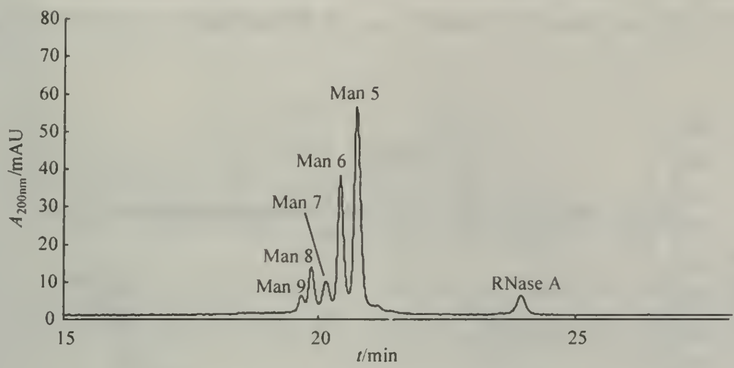


图 10.19　用极性反转方法对RNaseB的糖型进行CE-UV分析

（二）蛋白质高甘露糖型 N -糖基化和氧化情况的分析

对于蛋白质的 *N* -糖基化，可以考虑分为两个步骤：首先分析糖基化位点；然后再分析糖链结构。而对于高甘露糖型的 *N* -糖基化，可以很方便地通过比较性肽质谱方法来进行。*N* -糖苷酶F可以完全切断与肽链相连的糖链，其适宜pH值在8左右，正好可与胰蛋白酶或Lys-C蛋白酶联合使用；这二者的联合使用可以鉴定蛋白质的糖基化位点。α -甘露糖苷酶切断的是以 α -糖苷键连接的糖基团（在高甘露糖型的糖链中除了与GlcNAc相连的一个甘露糖外，其余的甘露糖均以 α -糖苷键彼此相连），其适宜的pH值在4左右，正好可与Glu-C蛋白酶联合使用；这二者的联合使用可以确定糖链的结构。

1. 糖基化位点和氧化情况的分析

TSL是在竹叶青蛇毒中发现的C型凝集素，它本身就具有结合半乳糖的活性。天然状态下表现为通过链间二硫键形成的多聚体。通过LC/MS方法对其还原态TSL及进一步用 *N* -糖苷酶F处理后的TSL进行的分子质量测定发现，还原态TSL中有分子质量不均一现象，且彼此之间相差约为162，含量最高的一种还原态TSL与用 *N* -糖苷酶F处理后的TSL之间的分子质量差距为1 703，正好符合 $GlcNAc_2Hex_8$ 的糖链结构。所以，还原态TSL中的糖型分布被归结为 $GlcNAc_2Hex_{5\sim8}$[78]。对比经 *N* -糖苷酶F处理前后的肽质谱图（未显示）发现，在别的酶解肽段出峰位置都基本保持不变的情况下，T1肽段经糖苷酶去除糖链后产生的T1′肽段的出峰位置与之发生了较大程度的改变（T1：1.3min；T1′：4.5min），这说明糖链的存在对该肽段的电泳迁移率影响很大。从TSL的氨基酸序列中可见，只有 *N* 端的T1肽段中含有符合特殊识别序列子的一个糖基化位点：Asn^5-Asp-Ser。另外一个有趣的现象是，在存在糖基化修饰的T1肽段中由于含有2个Met残基，所以还同时检测到了其中不同程度的Met残基氧化现象；在另一个含有Met残基的肽段T3中也同样检测到了它的氧化形式。经 *N* -糖苷酶F处理后的肽段T1′中同样也

检测到与 T1 中类似的 Met 残基氧化现象。

图 10.20a 显示了糖肽 T1 的糖型分布情况；图 10.20b 显示了 T1-GlcNAc$_2$Man$_8$型糖肽的氧化情况。由于在 T1-GlcNAc$_2$Man$_8$型糖肽（$S^{1}CCTN^{G}DSLPM^{10}NGM^{13}CYK^{16}$）中含有 Met^{10}和 Met^{13}两个 Met 残基，所以它产生了两种不同的氧化形式，即一个 Met 残基被氧化的形式（T1-GlcNAc$_2$Man$_8$-O）和两个 Met 残基同时被氧化的形式（T1-GlcNAc$_2$Man$_8$-2O）。

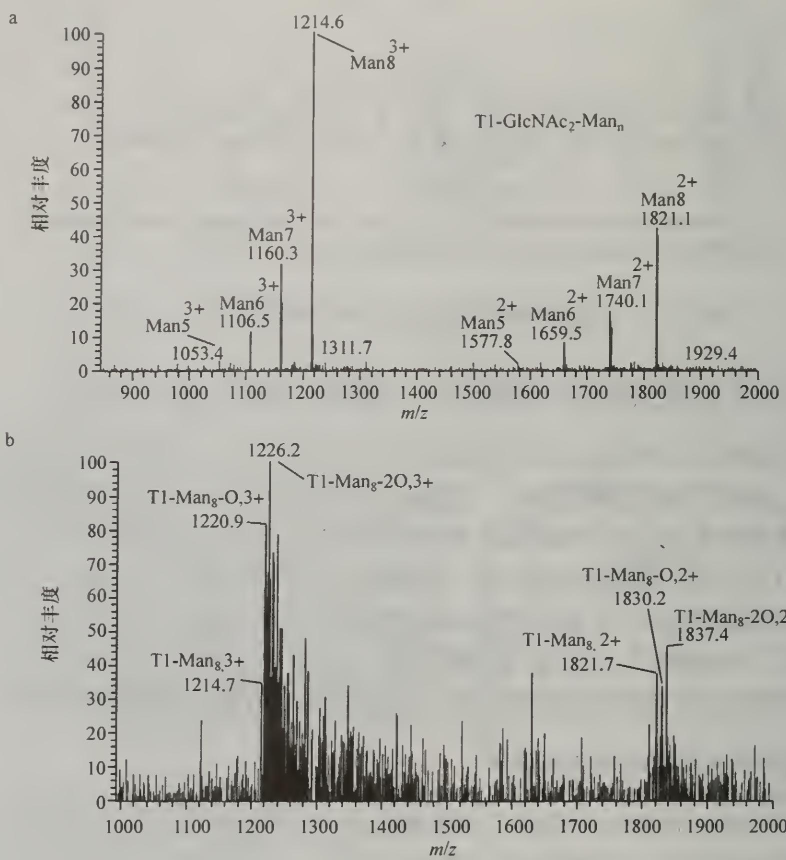


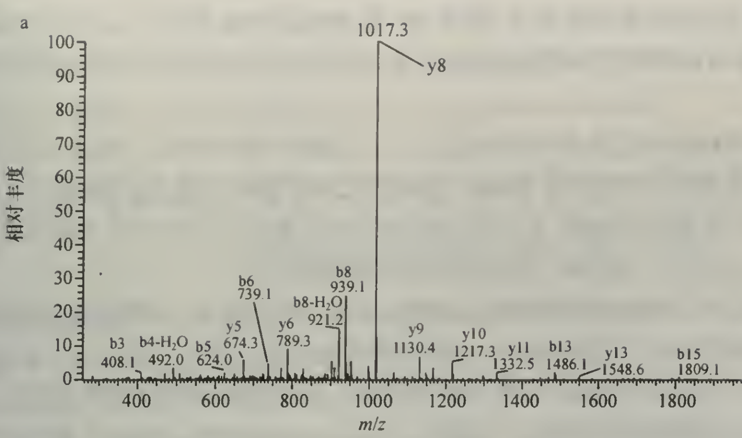
图 10.20　肽段 T1(a)及其氧化产物 T1-GlcNAc$_2$Man$_8$(b)的糖型分布质谱图

用 *N*-糖苷酶 F 对 TSL 的酶解产物进行处理后，糖肽 T1 的糖链被移去而转化为肽段 T1′，原来的 Asn^{5}残基转化为 Asp^{5}残基。尽管 T1′及其两种氧化形式并没有被分离开，但 CE/MS/MS 还是分别捕捉到了它们各自的 CID 质谱图。值得一提的是，通过 T1′-O（$S^{1}CCT\mathbf{D}^{5}DSLP\mathbf{M}^{10}NG\mathbf{M}^{13}CYK^{16}$）的 CID 质谱图可以清楚地看到，发生氧化的 Met 残基是 Met^{13}残基而不是 Met^{10}残基（图 10.21）。同样发生了 Met 残基氧化的还有肽段 T3（$T^{23}WEDAE\mathbf{M}^{29}FCR^{32}$）。

2. 糖链结构和氧化情况的分析

为了进一步验证 TSL 中糖链的高甘露糖型结构，需要用 α-甘露糖苷酶对它进行处

a



b	89.1	249.2	409.3	510.4	624.5	739.6	826.7	939.9	1037.0	1168.2	1282.3	1339.4	1486.6	1646.7	1809.9	1938.1	
	S	**C**	**C**	**T**	***D***	**D**	**S**	**L**	**P**	**M**	**N**	**G**	***M***	**C**	**Y**	**K**	
	1955.0	1867.9	1707.8	1547.7	1446.6	1332.5	1217.4	1130.4	1017.2	920.1	788.9	674.8	617.7	470.5	310.4	147.2	y

图 10.21　肽段 T1′-O 的 MS/MS 质谱图

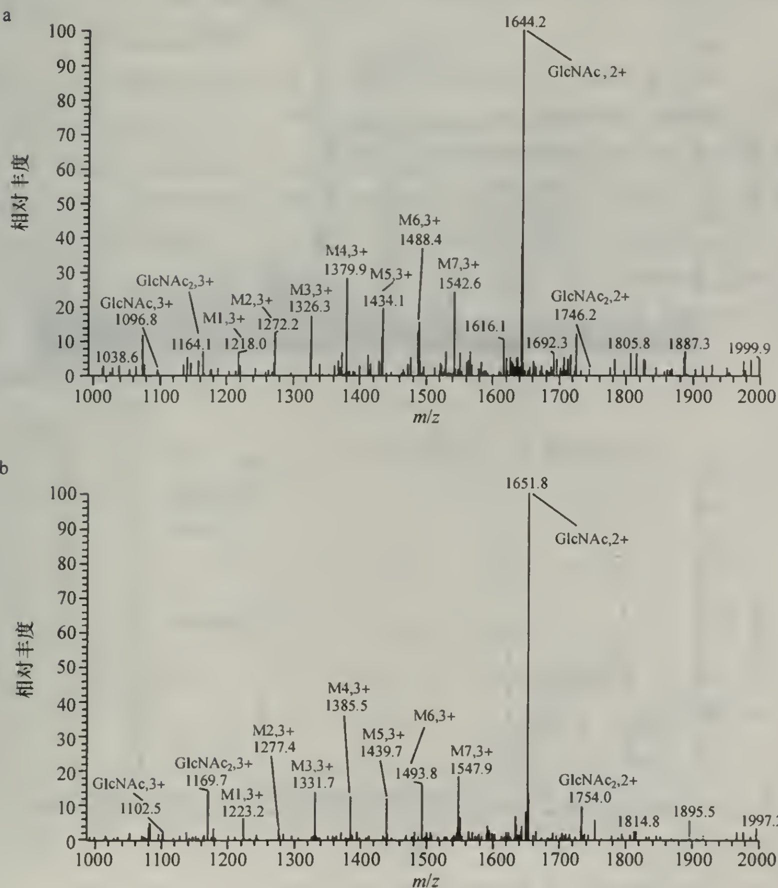


图 10.22　肽段 G1-GlcNAc$_2$Man$_8$(a)和 G1-GlcNAc$_2$Man$_8$ − O(b)的 MS/MS 质谱图

理。经过 6h 的 α-甘露糖苷酶处理后，糖肽 G1 的糖链结构由 $GlcNAc_2Man_8$ 转化为 $GlcNAc_2Man_3$ 结构，并且出峰位置由原来的比 G5 先出峰变为后来的比 G5 后出峰（图谱未显示）。

TSL 的 Met 残基的氧化现象同样也在其 Glu-C 酶解产物中存在。肽段 G1（$S^1CCT\mathbf{N}^5DSLP\mathbf{M}^{10}NG\mathbf{M}^{13}CYKIFDEPKTWE^{25}$）和 G4（$\mathbf{M}^{29}FCRKYKPGCHLASFHRLAE^{48}$）中均存在不同程度的 Met 残基氧化现象，由于这些肽段的氧化形式与它们本身的电泳迁移率差别很小，所以它们混和在一起出峰，造成相应峰的拖尾现象。

以 G1 为例来分析其分子中的糖基化和氧化情况。在未经 α-甘露糖苷酶处理前，糖肽 G1 的主要的糖链结构为 $GlcNAc_2Man_8$（糖肽表现为 1597.0^{3+}）。图 10.22 显示了 G1-$GlcNAc_2Man_8$ 和 G1-$GlcNAc_2Man_8$-O 的 CID 质谱图。可见这两种糖肽的断裂都主要发生在各种糖基团之间，因而得到了从 G1-$GlcNAc_2Man_7$～G1-$GlcNAc_1$ 之间一系列的完整的碎片离子，但糖链与肽链之间的共价键 G1↔GlcNAc 不发生断裂。这充分说明了糖链是 *N*-连接的高甘露糖型结构，而且这两种糖肽之间的差别仅仅在于其中一个糖肽中的 Met 残基发生了氧化。

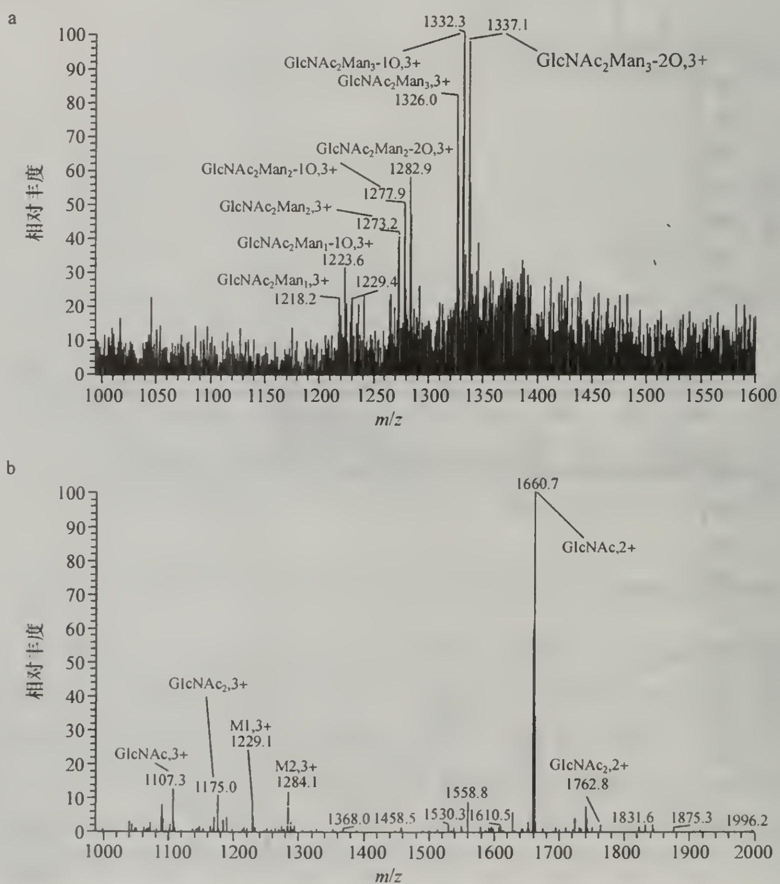


图 10.23　G1′-$GlcNAc_2Man_{1\sim3}$ 及其氧化产物的质谱图 a 及
G1′-$GlcNAc_2Man_3$-2O 的 MS/MS 质谱图 b

在经过 6h 的 α-甘露糖苷酶处理后，糖肽 G1 转化为糖肽 G1′(其中的糖链结构由 $GlcNAc_2Man_8$ 转化为 $GlcNAc_2Man_{1\sim3}$)。这也是高甘露糖型结构的一个佐证。图 10.23 中分别表示了糖肽 G1′的带电荷情况和 G1′-2O 的 MS/MS 情况。

五、蛋白质混合物的毛细管等电聚焦-质谱联用分析

(一) 标准蛋白混合物的 CIEF/MS 分析

首先用细胞色素 c、肌红蛋白和 β-乳球蛋白 A 的混合物试验了 CIEF/MS 系统的性能。样品为 1%(*m*/*V*)两性电解质 pH3.5～10，细胞色素 c 0.125mg/ml，肌红蛋白 0.03mg/ml，β-乳球蛋白 A 0.15mg/ml。聚焦电流开始时为 4.5μA，聚焦完成时变为 1.2μA。图 10.24 为 CIEF/MS 的 base peak 图谱。由图谱可以看到，这三种蛋白质的聚焦效果极好，出峰时间均仅在 1min 左右，这也是对等电点非常接近的混合物中的各种物质进行分辨的基础。另外，非常有趣的是，这三种蛋白质均被各自分辨出两种亚型(电荷分布和分子质量几乎完全一样)，且它们之间的等电点差异都很小。细胞色素 c 和肌红蛋白在等电聚焦条件下可以被分辨出两种亚型是有报道的[21]，而 β-乳球蛋白 A 的亚型则未见报道。

细胞色素 c 的已知等电点为 9.6[21]，肌红蛋白的两个已知等电点为 7.2 和 6.8[21]，β-乳球蛋白 A 的已知等电点为 5.1[79]。用它们各自的出峰时间对其等电点作线性回归，即可获得 CIEF/MS 的工作曲线。因为实验中所采用的条带移动方法是阳极电解质替换结合重力移动的方法，这种方法本身的移动速度并不是线性的，但在小的 pH 范围内可以保持良好的线性关系。根据在 pH 5～7 和 pH 6.5～10 范围的工作曲线，可以方便地得出细胞色素 c 的两种亚型的等电点分别为 9.60 和 9.55；肌红蛋白的两种亚型的等电点分别为 7.20 和 6.79；β-乳球蛋白 A 的两种亚型的等电点分别为 5.10 和 4.51。对这些亚型的分析情况总结在表 10.2 中。

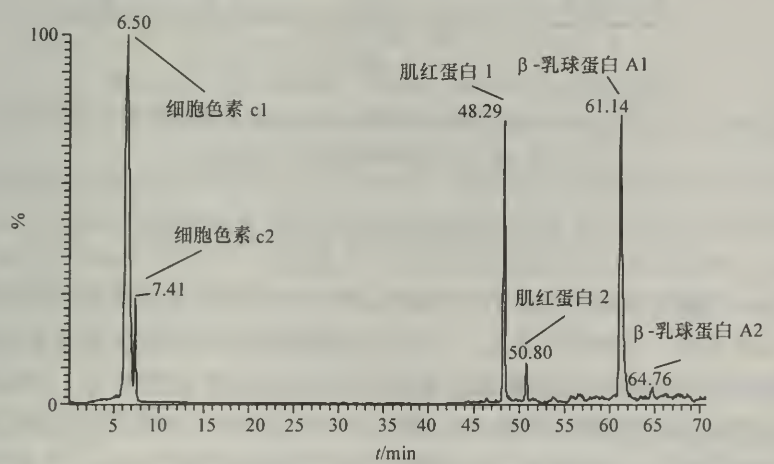


图 10.24　用 CIEF/ESI.MS 方法分析细胞色素 c、肌红蛋白和 β-乳球蛋白 A 的混合物(Pharmalyte 3.5～10)

表 10.2　用 CIEF-ESI-MS 方法分析细胞色素 c、肌红蛋白和 β-乳球蛋白 A 的混合物

蛋白质	保留时间/min	理论分子质量/Da	观测分子质量/Da	等电点	相对含量/%
细胞色素 c1	6.50	12359.3	12362.0	9.60	75.7
细胞色素 c2	7.41	12359.3	12363.0	9.55	24.3
肌红蛋白 1	48.29	16951.5	16952.0	7.20	80.7
肌红蛋白 2	50.80	16951.5	16953.0	6.79	19.3
β-乳球蛋白 A1	61.14	18367.3	18364.0	5.10	90.2
β-乳球蛋白 A2	64.76	18367.3	18364.0	4.51	9.8

用同样的方法分析了细胞色素 c，肌红蛋白和碳酸酐酶Ⅱ（pI = 5.9）的混合物（图 10.25）。根据 pH5.5～7.5 范围内的工作曲线，可以得出来自碳酸酐酶Ⅱ中的两种小分子杂质 M1 和 M2 的等电点分别为 6.15 和 5.55。另外，碳酸酐酶Ⅱ在实验中的 CIEF/ESI-MS条件下离子化非常好，在 m/z 1 100～2 000 范围内找到了它的 +15～ +26 离子。有趣的是，碳酸酐酶Ⅱ并没有亚型存在。

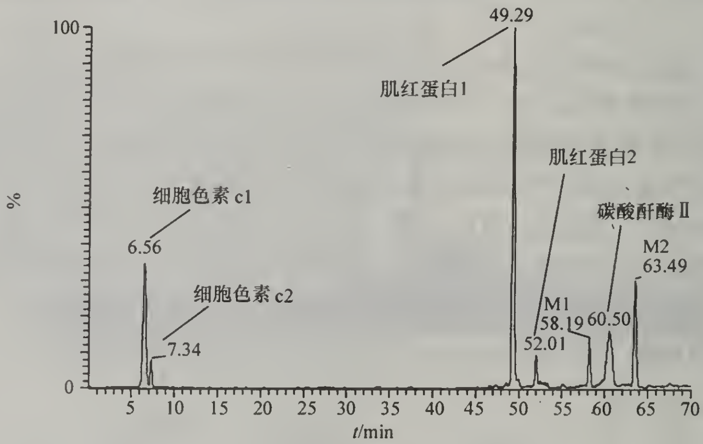


图 10.25　用 CIEF/MS 方法分析细胞色素 c、肌红蛋白和碳酸酐酶Ⅱ的混合物（M1 和 M2 是来自碳酸酐酶Ⅱ中的杂质）

（二）血红蛋白变种的 CIEF/MS 分析

图 10.26 中显示了分别用 pH 3.5～10 和 pH 5～8 的两性电解质对 HbS 和 HbA 混合物的 CIEF/MS 结果。尽管仅仅从 base peak 图谱看起来它们的分离度并不好，但是通过 EIM 方法对 HbS 和 HbA 进行分别的监测（选择 HbA 的 β 链的 + 12 离子，m/z = 1 323.4；对于 HbS，则选择其 β 链的 + 12 离子，m/z = 1 320.9）时则可发现，这主要是因为它们各自还有三种彼此之间等电点非常接近的亚型的缘故（图 10.26a），各种亚型之间的分离度是很好的。在用 pH 梯度更窄的两性电解质进行等电聚焦时（图 10.26b），各种亚型之间的分离度还可以得到提高。进一步对每一种亚型进行分析发现，HbS 的三种亚型的电荷分布情况和分子质量几乎完全一样；HbA 的三种亚型的电荷分布情况和分子质量

也几乎完全一样。

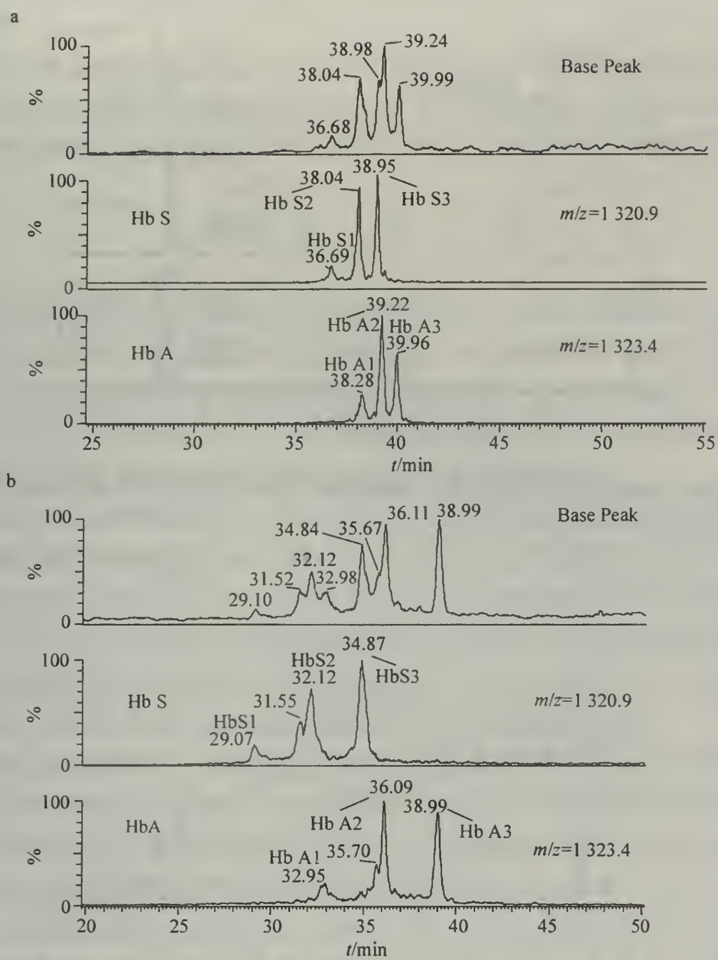


图 10.26 用 CIEF/MS 方法分析 HbS 和 HbA 的混合物

a. Pharmalyte 3.5～10; b. Pharmalyte 5～8

为了确定这些血红蛋白变种的亚型的等电点,需要加入标准蛋白与之共同进行等电聚焦,然后从工作曲线得出它们对应的等电点。图 10.27 是血红蛋白变种与细胞色素 c 及肌红蛋白的混合物的 CIEF/MS 图谱,图 10.28 是经线性回归得出的工作曲线。表 10.3 总结了这些血红蛋白变种的亚型的等电点及它们各自的相对含量。对于这些血红蛋白的亚型的分辨还未见任何文献报道。

本节的研究中采用 LPA 涂层的毛细管来消除电渗流带来的影响;用移动阳极电解质替换结合重力影响的方法来进行样品条带的移动;用活动式毛细管位置调整的方法来实现 CIEF 与 ESI-MS 的在线联用;用较长的毛细管来进行等电聚焦。所有这些技术和材料的使用所带来的结果就是:超高的等电点分辨率。对于其他的一般方法所无法分辨的等电点微小差异(如 $\Delta pI < 0.05$),可以比较容易地区分。如尽管 HbA2 和 HbA3 亚型之间的等电点差异只有 0.04,还是可以通过 CIEF/MS 得到近乎基线的分辨。

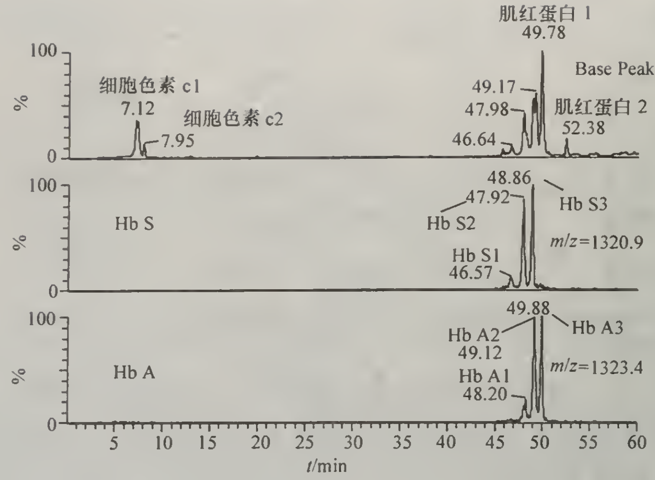


图 10.27　用 CIEF/MS 方法分析 HbS、HbA、细胞色素 c 和肌红蛋白的混合物(Pharmalyte 3.5～10)

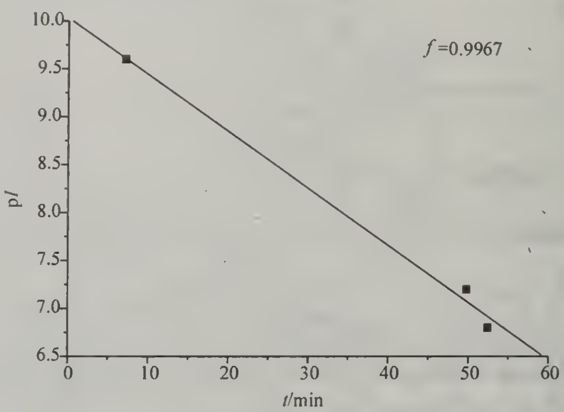


图 10.28　pH6.5～10 范围中的线性回归

表 10.3　用 CIEF-ESI-MS 方法分析 Hb S 和 Hb A 的混合物(Pharmalyte 3.5～10)

蛋白质	保留时间/min	理论分子质量/Da		观测分子质量/Da		等电点	相对含量/%
		α链	β链	α链	β链		
Hb S1	46.57	15126.4	15837.2	15127.0	15838.0	7.26	10.70
Hb S2	47.92	15126.4	15837.2	15128.0	15837.0	7.18	42.17
Hb S3	48.86	15126.4	15837.2	15128.0	15838.0	7.13	47.13
Hb A1	48.20	15126.4	15867.2	15128.0	15868.0	7.17	16.39
Hb A2	49.12	15126.4	15867.2	15128.0	15867.0	7.11	50.36
Hb A3	49.88	15126.4	15867.2	15128.0	15868.0	7.07	33.25

用于 CIEF/MS 分析的样品的制备比较简单，只需要将蛋白质溶于 1%的两性电解质水溶液中即可。当然，盐分较多的样品(如生理样品)必须经过脱盐的步骤，否则分辨率将有不同程度的下降。另外，蛋白质的浓度不能太大，一般宜<0.2mg/ml，因为 CIEF 方法

本身就有条带浓缩的效果，太多的样品不仅对分辨率有很大的影响，而且还非常容易造成毛细管的堵塞。

在超高分辨率条件下，采用 CIEF/MS 对蛋白质混合物进行分析还发现了另外一个非常有趣的现象，即多种蛋白质在 CIEF/MS 分析中都发现有不同的亚型存在。对于这种电荷分布情况和分子质量完全相同的亚型的存在，目前还没有一种确切的解释。

第五节　展　　望

CE/MS 联用技术从 1987 年问世以来，虽然时间不长，但无论在技术上还是应用上都已经获得了长足的进步。样品处理、CE 分离、接口、MS 检测等环节的每一次技术发展和创新，都会给生物学家和生化学家们带来惊喜，同时也进一步拓宽和加深了 CE/MS 在这些领域中的应用；而人们在应用 CE/MS 作各种研究的时候所发现的问题和提出的改进设想又可能对 CE/MS 技术本身的发展产生不可估量的推动作用。

CE 是目前为止分辨率最高的一维分离技术，而 MS 则是目前最灵敏的检测技术，而且 MS 还能提供丰富而精确的结构信息，这二者的结合所产生的巨大能量是可以想像的。在线预浓缩方法的产生解决了 CE 上样量小的问题；各种高效率接口的研制和应用极大程度地提高了 CE 向 MS 的离子转移效率；多种 CE 分离形式为不同物质的分析提供了极大的自由度；各种新型 MS 仪器的不断发展解决了 CE 的"后顾之忧"。所有的这一切令我们不难相信，CE/MS 将成为高精度和微量分析中不可或缺的技术。

质谱仪器的发展还将进一步朝着高灵敏度、高精确度、高分析速度和更强的多级质谱能力等方向发展。今后小型化的 IT-MS 将成为一般生物学实验室的常用仪器，因为其简单及操作易用性可以广为生物学家们所接受；而 FTICR-MS 及 Q-TOF-MS 等高端质谱仪器将在更加艰难和复杂的分析任务中显示它们的不凡实力。另外，ACE/MS 技术还具有很大的潜力，若能结合 MFD 设备作进一步的发展，利用多功能"CE 芯片"与 MS 联用进行高通量大规模自动化临床和药物分析将不再是梦想。而最近出现的高分辨率 CIEF/μESI-FTICR-MS技术使我们看到了 CE/MS 方法在蛋白质组分析方面的巨大潜力和优势，若相应的数据处理技术和数据库搜索的计算机算法能得到更快的发展，它必将受到蛋白质组研究者的更加广泛的关注和应用。不妨让我们作一个大胆的设想：若干年后，完全自动化的整合型 CE/MS 仪器将出现在普通的生物学实验室里。

（刘　韬）

参考文献

1 Dove A. Proteomics: translating genomics into products? Nat Biotechnol, 1999, 17: 233～236

2 Peck K, Stryer L, Glazer AN *et al*. Single-molecule fluorescence detection: autocorrelation criterion and experimental realization with phycoerythrin. Proc Natl Acad Sci USA, 1989, 86: 4087～4091

3 Figeys D, van Oostveen I, Ducret A *et al*. Protein identification by capillary zone electrophoresis/microelectrospray ionization-tandem mass spectrometry at the subfemtomole level. Anal Chem, 1996, 68: 1822～1828

4 Valaskovic GA, Kelleher NL, McLafferty FW. Attomole protein characterization by capillary electrophoresis-mass spectrometry. Science, 1996, 273: 1199～1202

5 Jorgenson JW, Lukacs KD. Capillary zone electrophoresis. Science, 1983, 222: 266～272
6 Issaq HJ. A decade of capillary electrophoresis. Electrophoresis, 2000, 21: 1921～1939
7 Novotny MV. Capillary electrophoresis. Curr Opin Biotechnol, 1996, 7: 29～34
8 Fenn JB, Mann M, Meng CK *et al*. Electrospray ionization for mass spectrometry of large biomolecules. Science, 1989, 246: 64～71
9 Karas M, Hillenkamp F. Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. Anal Chem, 1988, 60: 2299～2301
10 Whitehouse CM, Dreyer RN, Yamashita M *et al*. Electrospray interface for liquid chromatographs and mass spectrometers. Anal Chem, 1985, 57: 675～679
11 Olivares JA, Nguyen NT, Yonker CR *et al*. On-line mass spectrometric detection for capillary zone electrophoresis. Anal Chem, 1987, 59: 1230～1232
12 Moseley MA, Deterding LJ, Tomer KB *et al*. Nanoscale packed-capillary liquid chromatography coupled with mass spectrometry using a coaxial continuous-flow fast atom bombardment interface. Anal Chem, 1991, 63: 1467～1473
13 Smith RD, Udseth HR, Wahl JH *et al*. Capillary electrophoresis-mass spectrometry. Methods Enzymol, 1996, 271: 448～486
14 Ding J, Vouros P. Advances in CE/MS. Anal Chem, 1999, 71: 378A～385A
15 Figeys D, Aebersold R. High sensitivity analysis of proteins and peptides by capillary electrophoresis-tandem mass spectrometry: recent developments in technology and applications. Electrophoresis, 1998, 19: 885～892
16 Cai J, Henion J. Capillary electrophoresis-mass spectrometry. J Chromatogr A, 1995, 703: 667～692
17 Smith RD, Wahl JH, Goodlett DR *et al*. Capillary electrophoresis-mass spectrometry. Anal Chem, 1993, 65: 574A～584A
18 Smith RD, Olivares JA, Nguyen NT, *et al*. Capillary zone electrophoresis-mass spectrometry using an electrospray ionization interface. Anal Chem, 1988, 60: 436～441
19 Lee ED, Muck W, Henion JD *et al*. On-line capillary zone electrophoresis-ion spray tandem mass spectrometry for the determination of dynorphins. J Chromatogr, 1988, 458: 313～321
20 Severs JC, Smith RD. Characterization of the microdialysis junction interface for capillary electrophoresis/microelectrospray ionization mass spectrometry. Anal Chem, 1997, 69: 2154～2158
21 Yang L, Lee CS, Hofstadler SA *et al*. Characterization of microdialysis acidification for capillary isoelectric focusing-microelectrospray ionization mass spectrometry. Anal Chem, 1998, 70: 4945～4950
22 Kelly JF, Ramaley L, Thibault P. Capillary zone electrophoresis-electrospray mass spectrometry at submicroliter flow rates: pratical considerations and analytical performance. Anal Chem, 1997, 69: 51～60
23 Chang YZ, Her GR. Sheathless capillary electrophoresis/electrospray mass spectrometry using a carbon-coated fused-silica capillary. Anal Chem, 2000, 72: 626～630
24 Foret F, Muller O, Thorne J *et al*. Analysis of protein fractions by micropreparative capillary isoelectric focusing and matrix-assisted laser desorption time-of-flight mass spectrometry. J Chromatogr A, 1995, 716: 157～166
25 Walker KL, Chiu RW, Monnig CA *et al*. Off-line coupling of capillary electrophoresis and matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. Anal Chem, 1995, 67: 4197～4204
26 Zhang H, Caprioli RM. Capillary electrophoresis combined with matrix-assisted laser desorption/ionization mass spectrometry; continuous sample deposition on a matrix-precoated membrane target. J Mass Spectrom, 1996, 31: 1039～1046
27 Xue Q, Foret F, Dunayevskiy YM *et al*. Multichannel microchip electrospray mass spectrometry. Anal Chem, 1997, 69: 426～430
28 Wen J, Lin Y, Xiang F *et al*. Microfabricated isoelectric focusing device for direct electrospray ionization-mass spectrometry. Electrophoresis, 2000, 21: 191～197
29 Xiang F, Lin Y, Wen J *et al*. An integrated microfabricated device for dual microdialysis and on-line ESI-ion trap mass spectrometry for analysis of complex biological samples. Anal Chem, 1999, 71: 1485～1490

30 Li J, Kelly JF, Chernushevich I *et al*. Separation and identification of peptides from gel-isolated membrane proteins using a microfabricated device for combined capillary electrophoresis/nanoelectrospray mass spectrometry. Anal Chem, 2000, 72: 599~609

31 Foret F, Szoko E, Karger BL. On-column transient and coupled column isotachophoretic preconcentration of protein samples in capillary zone electrophoresis. J Chromatogr, 1992, 608: 3~12

32 Auriola S, Thibault P, Sadovskaya I *et al*. Enhancement of sample loadings for the analysis of oligosaccharides isolated from *Pseudomonas aeruginosa* using transient isotachophoresis and capillary zone electrophoresis-electrospray-mass spectrometry. Electrophoresis, 1998, 19: 2665~2676

33 Chien RL, Burgi DS. On-column sample concentration using field amplification in CZE. Anal Chem, 1992, 64: 489A~496A

34 Naylor S, Ji Q, Johnson KL *et al*. Enhanced sensitivity for sequence determination of major histocompatibility complex class I peptides by membrane preconcentration-capillary electrophoresis-microspray-tandem mass spectrometry. Electrophoresis, 1998, 19: 2207~2212

35 Figeys D, Ducret A, Yates JR 3rd *et al*. Protein identification by solid phase microextraction-capillary zone electrophoresis-microelectrospray-tandem mass spectrometry. Nat Biotechnol, 1996, 14: 1579~1583

36 Wahl JH, Gale DC, Smith RD. Sheathless capillary electrophoresis-electrospray ionization mass spectrometry using 10 mum I.D. capillaries: analyses of tryptic digests of cytochrome c. J Chromatogr A, 1994, 659: 217~222

37 Figeys D, Corthals GL, Gallis B *et al*. Datadependent modulation of solid-phase extraction capillary electrophoresis for the analysis of complex peptide and phosphopeptide mixtures by tandem mass spectrometry: application to endothelial nitric oxide synthase. Anal Chem, 1999, 71: 2279~2287

38 Liu T, Shao XX, Zeng R *et al*. Analysis of recombinant and modified proteins by capillary zone electrophoresis coupled with electrospray ionization tandem mass spectrometry. J Chromatogr A, 1999, 855: 695~707

39 Belov ME, Gorshkov MV, Udseth HR *et al*. Zeptomole-sensitivity electrospray ionization-Fourier transform ion cyclotron resonance mass spectrometry of proteins. Anal Chem, 2000, 72: 2271~2279

40 Ramjit HG, Kruppa GH, Spier JP *et al*. The significance of monoisotopic and carbon-13 isobars for the identification of a 19-component dodecapeptide library by positive ion electrospray Fourier transform ion cyclotron resonance mass spectrometry. Rapid Commun Mass Spectrom, 2000, 14: 1368~1376

41 Hofstadler SA, Wahl JH, Bruce JE *et al*. On-line capillary electrophoresis with Fourier transform ion cyclotron resonance mass spectrometry. J Am Chem Sco, 1993, 115: 6983~6984

42 Severs JC, Hofstadler SA, Zhao Z *et al*. The interface of capillary electrophoresis with high performance Fourier transform ion cyclotron resonance mass spectrometry for biomolecule characterization. Electrophoresis, 1996, 17: 1808~1817

43 Hofstadler SA, Severs JC, Smith RD *et al*. Analysis of single cells with capillary electrophoresis electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry. Rapid Commun Mass Spectrom, 1996, 10: 919~922

44 Wu JT, Qian MG, Li MX *et al*. Use of an ion trap storage/reflection time-of-flight mass spectrometer as a rapid and sensitive detector for capillary electrophoresis in protein digest analysis. Anal Chem, 1996,68: 3388~3396

45 Jin X, Kim J, Parus S *et al*. On-line capillary electrophoresis/microelectrospray ionization-tandem mass spectrometry using an ion trap storage/time-of-flight mass spectrometer with SWIFT technology. Anal Chem, 1999, 71: 3591~3597

46 Patterson SD, Aebersold R. Mass spectrometric approaches for the identification of gel-separated proteins. Electrophoresis, 1995, 16: 1791~1814

47 Katayama H, Ishihama Y, Oda Y *et al*. Electrophoretic mobility-assisted identification of proteins by nanoelectrospray capillary electrophoresis/mass spectrometry under methanolic conditions. Rapid Commun Mass Spectrom, 2000, 14: 1167~1178

48 Tong W, Link A, Eng JK *et al*. Identification of proteins in complexes by solid-phase microextraction/multistep elution/capillary electrophoresis/tandem mass spectrometry. Anal Chem, 1999, 71: 2270~2278

49 Cao P, Stults JT. Mapping the phosphorylation sites of proteins using on-line immobilized metal affinity chromatography/capillary electrophoresis/electrospray ionization multiple stage tandem mass spectrometry. Rapid Commun Mass Spectrom, 2000, 14: 1600～1606

50 Yeung B, Porter TJ, Vath JE. Direct isoform analysis of high-mannose-containing glycoproteins by on-line capillary electrophoresis electrospray mass spectrometry. Anal Chem, 1997, 69: 2510～2516

51 Boss HJ, Watson DB, Rush RS. Peptide capillary zone electrophoresis mass spectrometry of recombinant human erythropoietin: an evaluation of the analytical method. Electrophoresis, 1998, 19: 2654～2664

52 He T, Chandramouli N, Fu E *et al*. Analysis of reduced and oxidized forms of cytochrome c by capillary electrophoresis and capillary electrophoresis-mass spectrometry. Anal Biochem, 1999, 271: 189～192

53 Jensen PK, Harrata AK, Lee CS. Monitoring protein refolding induced by disulfide formation using capillary isoelectric focusing-electrospray ionization mass spectrometry. Anal Chem, 1998, 70: 2044～2049

54 Tang Q, Harrata AK, Lee CS. Two-dimensional analysis of recombinant *E. coli* proteins using capillary isoelectric focusing electrospray ionization mass spectrometry. Anal Chem, 1997, 69: 3177～3182

55 Zhang CX, Xiang F, Pasa-Tolic L *et al*. Stepwise mobilization of focused proteins in capillary isoelectric focusing mass spectrometry. Anal Chem, 2000, 72: 1462～1468

56 Jensen PK, Pasa-Tolic L, Anderson GA *et al*. Probing proteomes using capillary isoelectric focusing-electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry. Anal Chem, 1999, 71: 2076～2084

57 Jensen PK, Pasa-Tolic L, Peden KK *et al*. Mass spectrometric detection for capillary isoelectric focusing separations of complex protein mixtures. Electrophoresis, 2000, 21: 1372～1380

58 Veenstra TD, Martinovic S, Anderson GA *et al*. Proteome analysis using selective incorporation of isotopically labeled amino acids. J Am Soc Mass Spectrom, 2000, 11: 78～82

59 Shen Y, Berger SJ, Anderson GA *et al*. High-efficiency capillary isoelectric focusing of peptides. Anal Chem, 2000, 72: 2154～2159

60 Barroso MB, de Jong AP. Sheathless preconcentration-capillary zone electrophoresismass spectrometry applied to peptide analysis. J Am Soc Mass Spectrom, 1999, 10: 1271～1278

61 Cao P, Stults JT. Phosphopeptide analysis by on-line immobilized metal-ion affinity chromatography-capillary electrophoresis-electrospray ionization mass spectrometry. J Chromatogr A, 1999, 853: 225～235

62 Soudan B, Hennebicq S, Tetaert D *et al*. Capillary zone electrophoresis and MALDI-mass spectrometry for the monitoring of *in vitro* O-glycosylation of a threonine/serine-rich MUC5AC hexadecapeptide. J Chromatogr B, 1999, 729: 65～74

63 Forest E, Petillot Y, Gagnon J *et al*. Determination of the C-terminal form of an anemone toxin using capillary electrophoresis and mass spectrometry. Electrophoresis, 1996, 17: 962～964

64 Perkins JR, Paker CE, Tomer KB. The characterization of snake venoms using capillary electrophoresis in conjuction with electrospray mass spectrometry: black mambas. Electrophoresis, 1993, 14: 458～468

65 Ensing K, de Boer T, Schreuder N *et al*. Separation and identification of neuropeptide Y, two of its fragments and their degradation products using capillary electrophoresis-mass spectrometry. J Chromatogr B, 1999, 727: 53～61

66 Zhang H, Stoeckli M, Andren PE *et al*. Combining solid-phase preconcentration, capillary electrophoresis and off-line matrix-assisted laser desorption/ionization mass spectrometry: intracerebral metabolic processing of peptide E *in vivo*. J Mass Spectrom, 1999, 34: 377～383

67 Ivanov AR, Nazimov IV. Identification of unusual(modified and non-encoded)amino acid residues in peptides by combinations of high-performance liquid chromatography and high-performance capillary electrophoresis with matrix-assisted laser desorption ionization time-of-flight mass spectrometry. J Chromatogr A, 2000, 870: 255～269

68 Lynen F, Zhao Y, Becu C *et al*. Considerations concerning interaction characterization of oligopeptide mixtures with vancomycin using affinity capillary electrophoresis-electrospray mass spectrometry. Electrophoresis, 1999, 20: 2462～2474

69 Lyubarskaya YV, Dunayevskiy YM, Vouros P *et al*. Microscale epitope mapping by affinity capillary electrophoresis-

mass spectrometry. Anal Chem, 1997, 69: 3008～3014
70 Lyubarskaya YV, Carr SA, Dunnington D *et al*. Screening for high-affinity ligands to the Src SH2 domain using capillary isoelectric focusing-electrospray ionization ion trap mass spectrometry. Anal Chem, 1998, 70: 4761～4770
71 Tuffal G, Tuong A, Dhers C *et al*. Direct evidence of methylglucose lipopolysaccharides/palmitoyl-CoA noncovalent complexes by capillary electrophoresis-electrospray/mass spectrometry. Anal Chem, 1998, 70: 1853～1858
72 Gysler J, Mazereeuw M, Helk B *et al*. Utility of isotachophoresis-capillary zone electrophoresis, mass spectrometry and high-performance size-exclusion chromatography for monitoring of interleukin-6 dimer formation. J Chromatogr A, 1999, 841: 63～73
73 梁雪莲，王克夷．竹叶青蛇毒凝集素的分离纯化及特性研究．生物化学与生物物理学报，1993，25：515～521
74 Kelly JF, Locke SJ, Ramaley L *et al*. Development of electrophoretic conditions for the characterization of protein glycoforms by capillary electrophoresis-electrospray mass spectrometry. J Chromatogr A, 1996, 720: 409～427
75 Wiktorowicz JE, Colburn JC. Separation of cationic proteins via charge reversal in capillary electrophoresis. Electrophoresis, 1990, 11: 769～773
76 刘韬，曾嵘，邵晓霞等．毛细管区带电泳-串联质谱联用法鉴定多肽和蛋白质．生物化学和生物物理学报，1999，31：425～432
77 Wilm M, Neubauer G, Mann M. Parent ion scans of unseparated peptide mixtures. Anal Chem, 1996, 68: 527～533
78 Zeng R, Xu Q, Shao XX *et al*. Characterization and analysis of a novel glycoprotein from snake venom using liquid chromatography-electrospray mass spectrometry and Edman degradation. Eur J Biochem, 1999, 266: 352～358
79 Yang L, Tang Q, Harrata AK *et al*. Capillary isoelectric focusing-electrospray ionization mass spectrometry for transferrin glycoforms analysis. Anal Biochem, 1996, 243: 140～149

第十一章　毛细管电泳在蛋白质磷酸化和糖基化分析中的应用

蛋白质翻译后修饰是影响蛋白质结构和功能的重要因素，在生物化学途径的调控中起关键作用。而蛋白质磷酸化和糖基化为最常见的两种翻译后修饰，分析鉴定蛋白质磷酸化和糖基化是阐明他们如何调控生命活动的重要环节。目前已有许多方法手段分析鉴定这两种蛋白质修饰。在本章，我们只讨论毛细管电泳(CE)方法，并且重点介绍我们实验室的一些工作。

第一节　毛细管电泳分析多肽和蛋白质磷酸化

一、简　　介

蛋白质可逆磷酸化是最普遍最重要的一种蛋白质翻译后修饰(PTM)。在哺乳细胞中大约有1/3的蛋白质被认为是磷酸化修饰的[1]。蛋白质可逆磷酸化在胞外信息传递到核内，细胞生长、分裂、分化、代谢和癌症的产生等生命活动过程中起着关键的作用[1~3]，对众多蛋白质生物化学功能担负开/关调控责任，它是一种普遍的调控机制。当调控蛋白质或酶被特定的蛋白激酶在特定位点上磷酸化，该蛋白质或酶就被激活，或被抑制，或被标记上了其他调节因子作用的靶记号[4]。估计细胞内大约90%的磷酸化出现在丝氨酸残基，9.9%出现在苏氨酸残基，0.1%在酪氨酸残基[5, 6]。然而，目前仅有有限的磷酸化蛋白是已知的，而且许多涉及蛋白磷酸化激酶和磷酸酯酶的专一性，亦即它们的蛋白质底物仍不清楚。其主要原因之一是蛋白质磷酸化分析一直是一个技术挑战。虽然Edman降解氨基酸序列自动分析技术目前已十分成熟，但由于抽提氨基酸PTH衍生物的有机溶剂不能有效抽提亲水的磷酸氨基酸PTH衍生物，因此采用当前的蛋白质测序标准步骤不能分析磷酸氨基酸及其位点。而且，至今没有一个改进的自动测序方法对3种主要的磷酸氨基酸，即磷酸丝氨酸(pSer)、磷酸苏氨酸(pThr)和磷酸酪氨酸(pTyr)(图11.1)鉴定都通用。传统的方法是在体外用纯化的蛋白激酶标记放射性^{32}P于所研究的蛋白质或在体内作^{32}P代谢标记细胞，然后酶解标记的蛋白质，分离磷酸化肽，最后结合磷酸氨基酸分析和手工Edman降解氨基酸序列分析，在每一测序循环检测放射性丢失，鉴定磷酸氨基酸残基和位点。Meyer等用气相蛋白质顺序仪固相非放射法分析磷酸酪氨酸[7]。但是由于磷酸酪氨酸PTH衍生物在反相HPLC柱上出峰很宽，且峰只比基线稍高一些，难以得到肯定的鉴定。为此，他们用毛细管电泳进一步验证序列仪反相柱流出液中是否存在磷酸酪氨酸PTH衍生物。我们报道了一种有效的毛细管电泳方法[8]，可以在低pmol范围同时非放射分析3种常见PTH-磷酸氨基酸，进而可分析蛋白质或多肽中的磷酸氨基酸。还可用此方法侦测磷酸化蛋白酶解产物中磷酸化肽段。

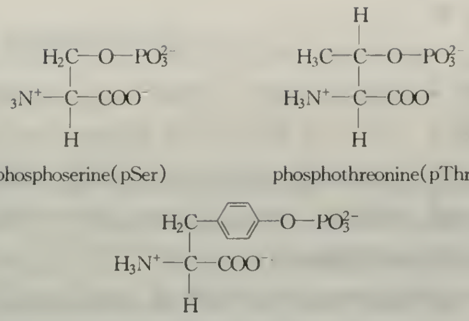


图 11.1　常见 3 种磷酸氨基酸结构

二、材　　料

①*L*-磷酸丝氨酸、*L*-磷酸苏氨酸、*L*-磷酸酪氨酸、牛 β-酪蛋白、卵黄高磷蛋白。顺序级苯异硫氰酸酯(PITC)、乙腈、三氟乙酸和乙酸乙酯为 ABI(Applied Biosystems)公司的产品。

②七肽 Leu-Arg-Arg-Ala-Ser-Leu-Gly(Kemptide)、六肽 Lys-Arg-Thr-Leu-Arg-Arg(hexapeptide)、蛋白激酶 A 催化亚基和蛋白激酶 C 均为 Sigma 公司产品。含磷酸酪氨酸残基的合成肽(phosphorylated hendecapeptide)Glu-Asp-Asn-Glu-Tyr(PO_3H)-Thr-Ala-Arg-Gln-Gly-Ala 由意大利 Lorenzo A. Pinna 教授赠送。

③丝氨酸磷酸化肽 Leu-Arg-Arg-Ala-Ser(PO_3H)-Leu-Gly(kemptide-P)和苏氨酸磷酸化肽 Lys-Arg-Thr(PO_3H)-Leu-Arg-Arg(hexapeptide-P)为本实验室分别用蛋白激酶 A 催化亚基和蛋白激酶 C 催化磷酸化相应底物肽,然后用 HPLC 纯化而得。

三、方　　法

(1) 部分酸水解

水解和衍生均在 Waters Pico. Tag Work Station 上进行。蛋白或多肽样品先吸入 5mm×50mm 硼硅酸玻璃管底部,然后将玻璃管放入水解反应瓶,真空抽干。吸 250μl 5.7mol/L 恒沸盐酸(含 1%重蒸苯酚)于反应瓶底部。小心抽气至盐酸溶液开始冒气泡,关闭真空阀,打开高纯氮气阀充氮气。如此重复两次后,密封反应瓶,放入 110℃ 反应釜中,水解 1.5h。取出冷至室温,高真空下抽干。

(2) PTH 衍生反应

磷酸氨基酸 PTH 衍生反应基本按照 Meyer 方法[7]。

条件溶液 1:以体积比 2:2:1 混合乙醇、三乙胺和水;

条件溶液 2:以体积比 8:1:1 混合乙醇、三乙胺和水;

衍生溶液:以体积比 7:1:1:1 混合乙醇、三乙胺、水和 PITC。

衍生步骤:

①加 10μl 条件溶液 1 于部分酸水解后抽干的样品管,稍后高真空下抽干(真空度小

于 0.04mbar)。重复此步骤一次。

②加 10μl 条件溶液 2，充高纯氮。密封反应瓶后，室温放置 5min。

③加 10μl 衍生溶液，充高纯氮，50℃ 下保温 5min。高真空下抽干(真空度小于 0.04mbar)。

④加 50μl 乙酸乙酯，轻轻旋转样品管，然后倒出乙酸乙酯。稍微抽干后，重复该洗涤步骤。

⑤加 50μl 20%三氟乙酸，充高纯氮。密封反应瓶后，50℃下保温 30min。高真空下抽干。

(3) 毛细管电泳分析 PTH-磷酸氨基酸

将 PTH 衍生物溶于 20%乙腈，0.1%三氟乙酸，25mmol/L NaCl 样品溶液中。毛细管为熔融石英管，总长 72cm，有效长度(进样端至检测器)50cm，内径 50μm；电泳缓冲液：25mmol/L NaH_2PO_4-H_3PO_4，pH2.5，15mmol/L NaCl；电压：-22kV；电压极性：负；检测波长：261nm；毛细管恒温温度：(30±0.2)℃。

四、结果和讨论

1. 毛细管电泳分离 PTH-磷酸氨基酸

在 pH2.5 的电泳缓冲溶液中，所有这 3 种 PTH-磷酸氨基酸都带负电荷，而其他 PTH-氨基酸带正电荷或不带电荷。当施加负相电压时，只有 PTH-磷酸氨基酸逆电渗流方向迁移，通过检测器而被检测，而其他 PTH-氨基酸顺电渗流方向迁移，不会通过检测器(图 11.2)。因此，这种磷酸氨基酸分析方法的选择性较高。图 11.3 为三种常见磷酸氨基酸(pSer、pThr 和 pTyr)的 PTH 衍生物毛细管电泳分离图。

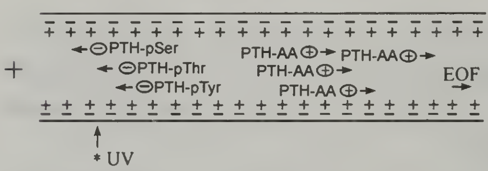


图 11.2 毛细管电泳分析 PTH-磷酸氨基酸示意图

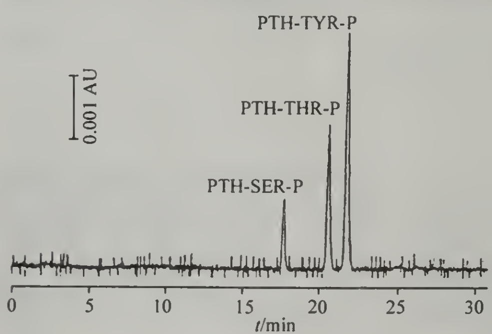


图 11.3 毛细管电泳分离 PTH-磷酸氨基酸

磷酸丝氨酸、磷酸苏氨酸和磷酸酪氨酸各 2nmol 混合，用 PITC 衍生。衍生物溶于 20μl 样品溶解液中，真空进样 2s(约 7nl)。条件：熔融石英毛细管，总长 72cm，有效长度 50cm，内径 50μm；电泳缓冲液：25mmol/L NaH_2PO_4-H_3PO_4，15mmol/L NaCl，pH2.5；电泳电压：-22kV；检测波长：261nm

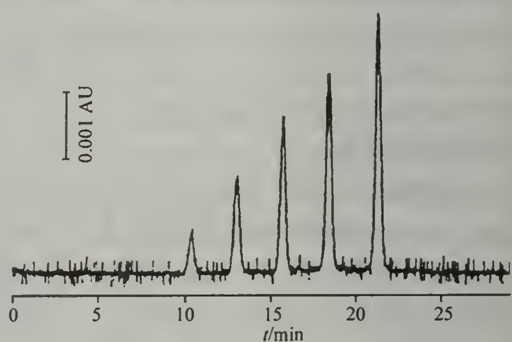


图 11.4 间隙进样技术同时分析不同浓度 PTH-磷酸酪氨酸

磷酸酪氨酸 1，2，3，4 和 5nmol，分别用 PITC 衍生。衍生物分别溶于 40μl 样品溶解液中。每一样品真空进样 2s。进样方法：进样第一号样品→电泳 2min→停止电泳→进样第二号样品→电泳 2min→停止电泳→……→进样第五号样品→电泳并记录电泳图。其余条件同图 11.3

2. 定量分析磷酸氨基酸的线性关系

利用毛细管电泳间隙进样技术[9]可以一次同时电泳同一种磷酸氨基酸但不同浓度的 PTH 衍生物，这样既快速又准确地获得定量关系。图 11.4 是采用该间隙进样技术，同时电泳 5 个浓度的磷酸酪氨酸 PTH 衍生物的电泳图谱。上述 3 种磷酸氨基酸的浓度(25～250μmol/L 范围)与其 PTH 衍生物毛细管电泳 UV 吸收峰高一次线性回归曲线的线性相关系数均大于 0.992($n=6$)。磷酸酪氨酸、磷酸苏氨酸和磷酸丝氨酸的质量检测限分别为 170fmol、350fmol 和 700fmol(指吸入毛细管的绝对量)。

3. 多肽及蛋白质中磷酸化氨基酸的分析

当 2nmol Kemptide-P 经 1.5h 酸水解，用 PITC 衍生后，CE 分析证实 Kemptide-P 为丝氨酸磷酸化肽(图 11.5a)。图 11.5b 和图 11.5c 分别是对磷酸化肽 hexapeptide-P 和磷酸化十一肽分析结果，证实他们分别含磷酸苏氨酸残基和磷酸酪氨酸残基。牛 β-酪蛋白是天然的磷酸化蛋白，有 209 个氨基酸残基，其中有 5 个丝氨酸残基被磷酸化。约 1nmol 牛 β-酪蛋白直接经酸水解 1.5h 衍生。毛细管电泳分析证实它是丝氨酸磷酸化蛋白(图 11.5d)。鸡卵黄高磷蛋白有 271 个氨基酸残基，是磷酸化程度最高的天然磷酸化蛋白之一，10%的残基被磷酸化。毛细管电泳分析显示它含磷酸丝氨酸残基(图 11.5e)。

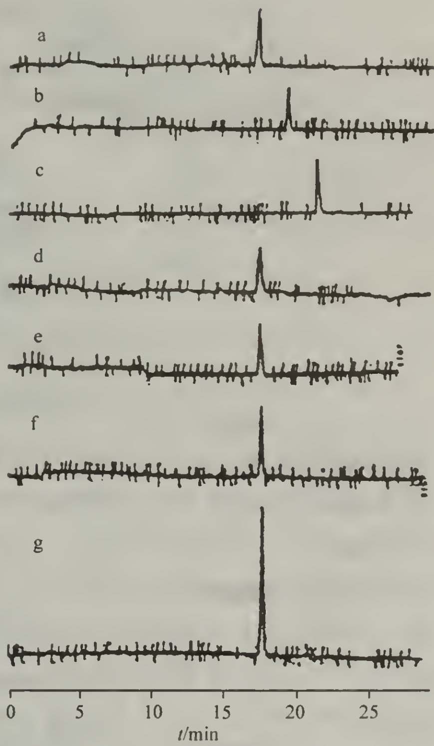


图 11.5　CE 分析肽或蛋白质中磷酸氨基酸
肽或蛋白质经部分酸水解 1.5h。水解产物用 PITC 衍生。PTH 衍生物溶解于 5μl 样品溶解液中(除图 11.5E 外)。a. 2nmol 含 Kemptide-P，进样 5s；b. 2nmol hexapeptide-P，进样 4s；c. 1nmol 含磷酸酪氨酸十一肽，进样 1s；d. 1nmol β-酪蛋白，进样 2s；e. 0.5nmol 卵黄高磷蛋白，衍生物溶于 100μl 样品溶解液中，进样 2s；f. 100μmol/L PTH-磷酸丝氨酸标准，进样 2s；g. 样品 E 和 100μmol/L PTH-磷酸丝氨酸标准各进样 2s。其余条件同图 11.3

水解条件是分析磷酸化肽和磷酸化蛋白的最关键因素。在给定的水解条件下，某一磷酸氨基酸的水解回收率对于不同的肽或蛋白质是不同的。磷酸丝氨酸的回收率尤其依赖于肽的序列[10]，而且磷酸丝氨酸对强酸的敏感性(指发生 β-消除或磷酸酯水解)高于磷酸苏氨酸和磷酸酪氨酸，这使得对其分析尤为困难。对大多数磷酸化肽来说，通常水解 1～3h 可得到较好的磷酸化氨基酸回收率。对于特定的磷酸化蛋白或磷酸化肽，欲获得最大的磷酸氨基酸回收率，需要优化水解条件(温度、时间和酸浓度)。分析蛋白质中磷酸氨基酸，也可先用测序级的高纯度蛋白水解酶将待分析的蛋白质降解成肽，然后对肽混合物进行酸水解。这样磷酸氨基酸回收率有可能高于直接酸水解待分析的蛋白质。这时，为排除蛋白水解酶样品含有的杂质蛋白是磷酸化蛋白，最好做对照实验。另外，所用的蛋白水解酶还应当不含磷酸酯酶杂质蛋白。

第二节 毛细管电泳分析单糖和寡糖

一、简　　介

治疗疾病用的一系列重组糖蛋白如重组促红细胞生成素(rEPO)、重组组织纤溶酶原激活物(rt-PA)等在临床的成功应用,以及其他药用重组糖蛋白的研发,日益激发人们对糖蛋白中糖部分的性质的关注和兴趣。愈来愈多的证据表明这些寡糖链的组成和结构影响糖蛋白的特性和生物活性,糖基化在细胞识别和免疫应答中起重要作用。图 11.6 给出在糖蛋白和蛋白聚糖中常出现的单糖及其结构。糖蛋白中的寡糖链就是这些单糖通过

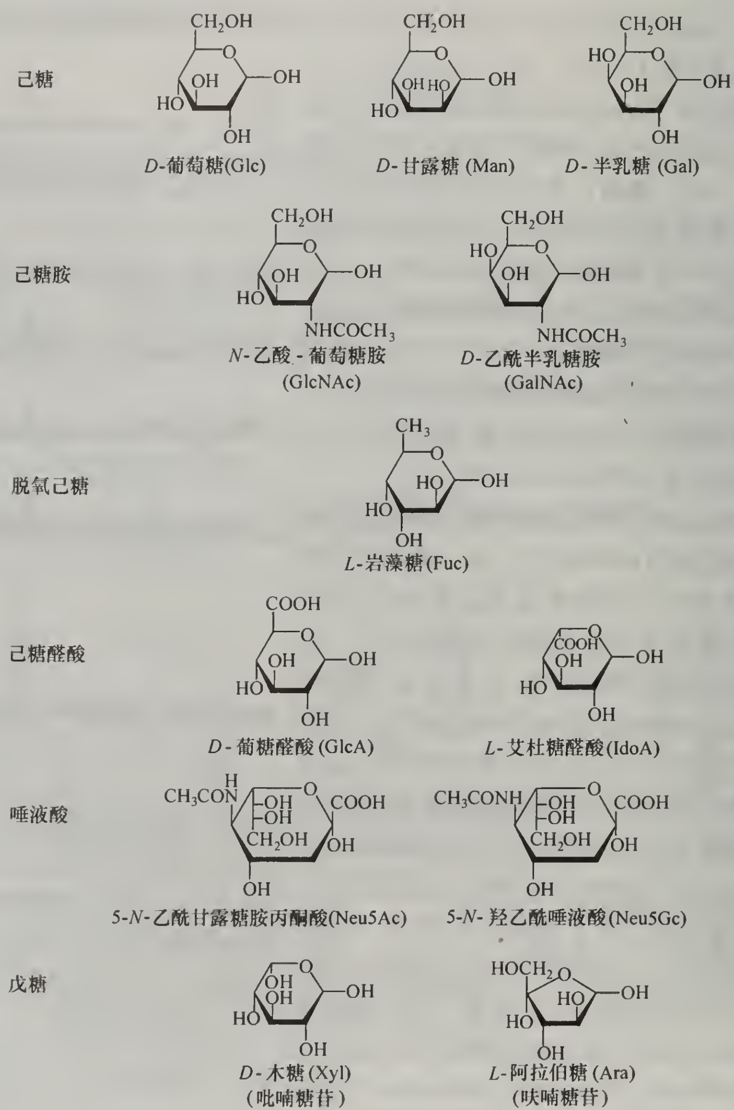


图 11.6　糖蛋白或蛋白聚糖中的常见单糖及其结构

α-或β-糖苷键以及不同的支链形式连接成结构复杂的寡聚物。糖链经 N-糖苷键与多肽的天冬酰氨残基侧链连接，或经 O-糖苷键与多肽的氨基酸残基的羟基连接，如通常与丝氨酸或苏氨酸残基侧链上羟基连接，有时与羟脯氨酸或羟赖氨酸残基的羟基相连。N-糖基修饰几乎总出现在 Asn-X-Ser/Thr 三肽序列(这里 X 为除 Pro 或 Asp 外的任何氨基酸残基)[11]；相反，对于 O-糖基修饰，目前还没鉴定到类似的糖基转移酶底物识别序列。一般 N-连接寡糖分为三大类，即高甘露糖型、复杂型、杂合型[11]。图 11.7 对每一类 N-连接寡糖糖型列举一个代表性例子。从图 11.7 可见，这三类 N-连接糖型都有一个相同的保守的由五元糖构成的核心结构，即 Manα1→3(Manα1→6)Manβ1→4GlcNAcβ1→4GlcNAc。不同 N-连接寡糖链就是在这五元糖核心结构基础上进行不同的分支和延伸而形成的。而与 N-连接寡糖链相比，O-连接寡糖链短，但结构多变化，没有一致的核心结构为 O-连接寡糖共同享有[12]，这使 O-连接寡糖链结构的分析更加困难。

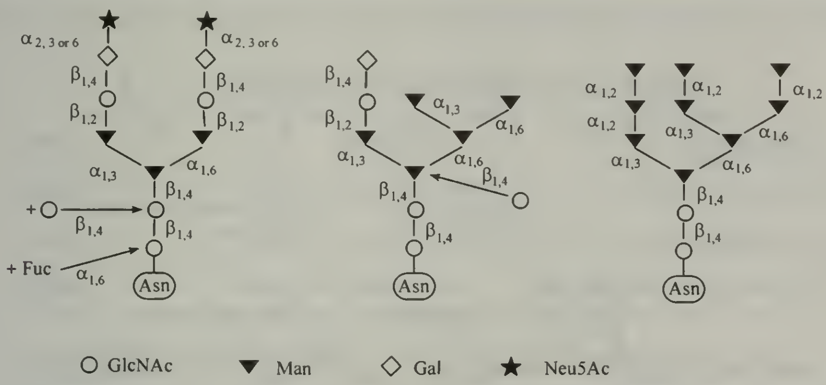


图 11.7　糖蛋白中 N-连接寡糖分类

糖分子有两个特性。一是缺乏常规紫外吸收基团和荧光基团，只在低紫外波长(如 195nm)范围才有光吸收；二是除酸性糖(如唾液酸、葡萄糖醛酸、半乳糖醛酸和磺酸化修饰糖等)外，糖分子因其羟基解离常数极小($10^{-11}\sim10^{-13}$)，是非常弱的酸，只在强碱条件下(pH>12)由于羟基的微弱电离[13]而带部分电荷。另外，糖分子能与硼酸盐形成带负电荷的复合物。基于第一个特性，检测糖分子的方法分衍生化法和非衍生化法两类。前者包括紫外[14~16]、荧光和激光诱导荧光检测法[17, 18]；后者包括直接紫外法[19, 20]、间接紫外法[21]、间接荧光法[22]、化学检测法[23]和折射法等。基于第二个特性，可以用硼酸盐或高碱性(如 NaOH)缓冲液为毛细管电泳液，使不同结构的糖分子具有不同的电迁移率而得以分离。当然，若衍生试剂本身带电荷，那么衍生化不仅为糖分子提供紫外吸收基团或荧光基团，同时还提供电迁移必要的电荷，因而电泳缓冲液不必局限于硼酸盐或强碱性(如 NaOH)溶液。

目前，有三种化学衍生方法衍生糖化合物。最常见的是还原胺化法[14~16]，以及 1-苯基-3-甲基-5-吡唑酮(1-phenyl-3-methyl-5-pyrazolone, PMP)缩合法[24]和羧基化糖与胺缩合法[25]。图 11.8 列出了常用的糖分子衍生剂，紫外检测波长，激光诱导荧光检测的激发波长和发射波长以及有关的部分参考文献。图 11.9 显示了还原单糖与 AMAC 还原胺化的反应历程。还原糖的醛基与衍生试剂的氨基反应形成 Schiff 碱。Schiff 碱的形成

是动力学快速平衡反应，但热力学不利。因此用硼氰氢化钠不断地还原平衡反应产物 Schiff 碱，迫使衍生反应进行完全，形成稳定的二级胺[37]。还原反应可能是该衍生反应的总限速步骤[37]。为使反应完全，并避免二级胺进一步与醛基反应，反应体系中应有至少 5 倍过量的衍生试剂存在[38]。

从糖蛋白释放单糖的常用方法是酸水解。也可用一系列的针对某一单糖残基及其连接于糖链上的方式(如 α-或 β-糖苷键，1,3 或 1,4 连接，等等)专一的外切糖苷水解酶从糖链的非还原端逐步切除单糖[39, 40]。酸水解主要用于糖蛋白的单糖组成分析，而外切糖苷水解酶法主要用于某一寡糖的糖序列(结构)分析。释放完整糖链的方法有碱降解法(β-消除反应)、肼解法和内切糖苷水解酶酶切法[12](用 *N*-糖苷水解酶切除 *N*-连接寡糖链，*O*-糖苷水解酶切除 *O*-连接寡糖链)。如果选用肼解法，需要对肼解的产物重乙酰化。

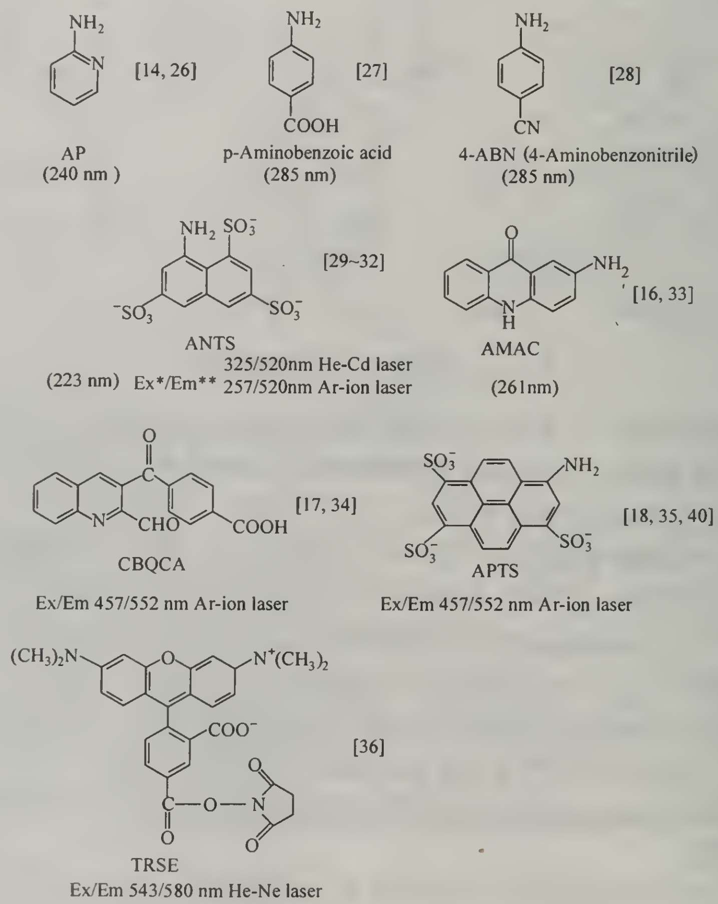


图 11.8 常见糖分子衍生剂

* Ex = Excitation; * * Em = Emission

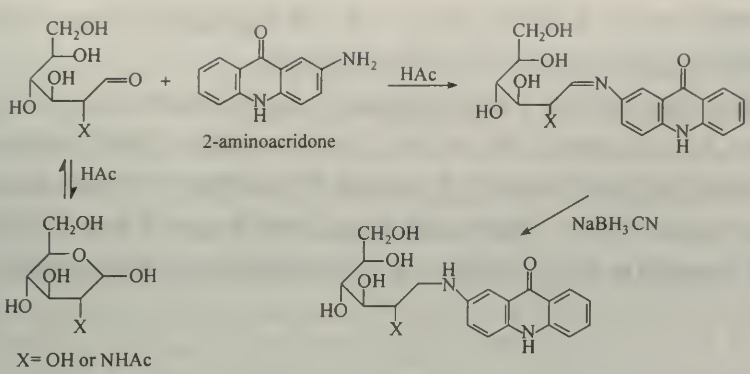


图 11.9　还原单糖与 AMAC 还原胺化反应历程示意图

表 11.1 列出分别从完整糖蛋白和纯化的寡糖酸水解释放单糖的通常条件[35, 41]。

表 11.1　从完整糖蛋白和纯化的寡糖酸水解释放单糖的条件

单糖	完整糖蛋白	纯化的寡糖
唾液酸	0.1mol/L TFA, 80℃, 0.5～1h	0.1mol/L TFA, 80℃, 0.5～1h
中性糖	2mol/L TFA, 100℃, 5h	2mol/L TFA, 100℃, 1～3h
氨基糖	4mol/L HCl, 100℃, 3h	4mol/L HCl, 100℃, 1h

对于乙酰氨基糖的分析,酸水解后,为了重新得到 *N*-乙酰氨基葡萄糖和 *N*-乙酰氨基半乳糖,也需要进行重乙酰化反应。进行稍大规模(对于几毫克到数十毫克糖蛋白分析)重乙酰化反应及其反应后样品脱盐可参考本文和文献[17];而进行小规模(对于几微克到数十微克糖蛋白分析)重乙酰化反应时[18],用 5μl 25mmol/L (pH9.5)碳酸盐溶液将

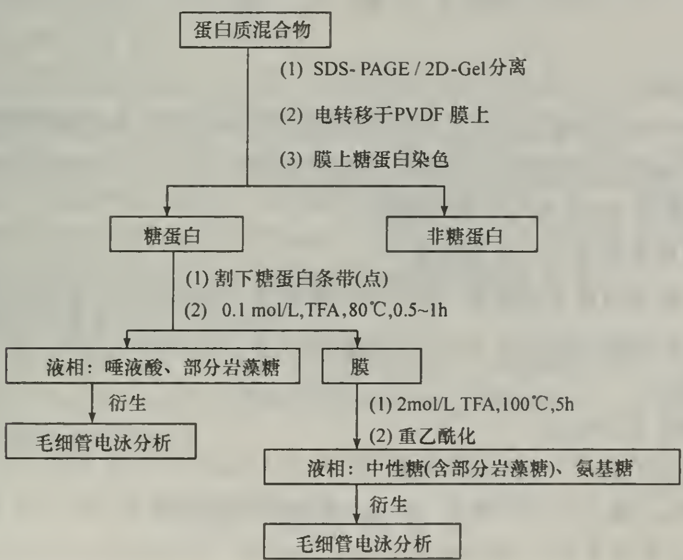


图 11.10　毛细管电泳分析 PVDF 膜上糖蛋白中单糖成分流程图

酸水解后单糖混合物冻干重新溶解，加 2μl 无水乙酸酐，室温反应 30min 后，冻干，再经衍生反应后，直接用毛细管电泳分析，就可得到较好结果。

如果需要分析混合物中某一糖蛋白的糖组成，可用 SDS-PAGE 电泳（对于更复杂的蛋白质混合物，如组织或细胞的蛋白抽提物，用 2-D Gel 电泳）分离纯化蛋白，电转移分离的蛋白质于聚偏二氟乙烯（PVDF）膜上，运用糖蛋白染色法[40~42]检测膜上糖蛋白条带（点）。从平行实验得到的另一张经考马斯亮蓝染色的 PVDF 膜上割下感兴趣的糖蛋白条带（点），酸水解释放糖蛋白中的单糖。最后，用毛细管电泳分析单糖衍生物。实验流程见图 11.10。

二、材　　料

①单糖：*N*－乙酰氨基半乳糖（GalNAc），木糖（Xyl），核糖（Rib），*N*－乙酰氨基葡萄糖（GlcNAc），葡萄糖（Glc），甘露糖（Man），阿拉伯糖（Ara），岩藻糖（Fuc），半乳糖（Gal），葡萄糖醛酸（GlcUA），半乳糖醛酸（GalUA）和 5－*N*－乙酰氨基唾液酸（Neu5Ac）。

②聚糖：葡聚糖（Dextran，平均相对分子质量为 523 000）。

③衍生剂：2－氨基 啶酮（2-Aminoacridone, AMAC），8－氨基萘基－1，3，6－三磺酸盐（8-aminonaphthalene-1,3,6-trisulfonate, ANTS）。

④糖蛋白：火鸡卵白蛋白（Sigma 产品 A5015），核糖核酸酶 B，α_1－酸性糖蛋白（α_1-AGP），人尿胰蛋白酶抑制剂（hu-UTI），重组人促红细胞生成素（rhu-EPO）。

⑤江浙蝮蛇毒蛋白：江浙蝮蛇毒粗品经硫酸铵反透析沉淀得到蛇毒蛋白混合物。

⑥其他试剂：四氢呋喃（THF），二甲亚砜（DMSO），三氟乙酸（TFA），无水乙酸酐，硼氰氢化钠，硼酸，硼砂，冰乙酸。

三、方　　法

(1) 无水肼的制备

30g 80％肼，150 g 甲苯及 150 g 氧化钙在圆底烧瓶内混合，室温放置过夜。混合物在装有氢氧化钠干燥管的回流装置回流 1h 后，蒸馏，收集 93～94℃馏份。甲苯与无水肼分层，收集下层无水肼，密封，置 4℃暗处保存。

(2) 糖蛋白肼解释放完整 *N*－连接寡糖

将 1～2mg 糖蛋白置于安瓿瓶，加入 1ml 新制无水肼溶解，充高纯氮，封管，100℃反应 8～10h。转移肼解产物于 1.5ml 的 Eppendorf 管，离心冷冻干燥。加 400μl 甲苯，振荡，离心冷冻干燥。

(3) 酸水解寡糖及重乙酰化

转移肼解后的样品于一洁净安瓿瓶，溶于 1ml 2mol/L 三氟乙酸，充高纯氮，封管，100℃水解 1～3h。离心冷冻干燥后，加 1ml 饱和碳酸氢钠溶解，加 10μl 无水乙酸酐，混匀，反应 10min。重复 3 次后，将反应物过 Bio-Rad RG-501-X8（H^+）柱，用 5 倍床体积的 Milli-Q 水洗脱柱。合并流出液和洗脱液，离心冷冻干燥。

(4) 酸水解释放糖蛋白中的唾液酸

分别取 1nmol α_1-AGP, hu-UTI 和 rhu-EPO 于 3 只 0.5ml Eppendorf 管, 离心冷冻干燥, 各加 100μl 0.1mol/L TFA。在 80℃水浴保温反应 0.5 至 1.0h, 离心冷冻干燥。

(5) 酸水解释放 PVDF 膜上糖蛋白中单糖

将 SDS-PAGE 分离并电转移至 PVDF 膜上的蛋白用考马斯亮蓝(0.1% R-250, 50%甲醇)染色 3min, 然后用 50%甲醇, 10%乙酸溶液脱色, 水洗 3～4 次后, 晾干。用洁净的剪刀剪下糖蛋白条带, 并将该条带剪成 2～3mm 宽, 长 0.5～0.8cm 小条带, 放入一只 0.5ml Eppendorf 管, 加 200μl 0.1mol/L TFA, 密封, 置 80℃水浴反应 40min。将水解液转移至另一只 0.5ml Eppendorf 管, 再用 200μl 水分两次浸洗 PVDF 膜, 合并水解液与浸洗液, 离心冷冻干燥, 放置 -20℃待衍生分析唾液酸。将 PVDF 膜小条带转移至一洁净 5cm×5mm 硼硅酸盐玻璃水解管底部, 加 200μl 2mol/L TFA。然后将水解管放入水解瓶, 密闭水解瓶, 置 100℃反应釜, 水解 4h。冷至室温后, 吸出水解液入 0.5ml Eppendorf 管, 用 200μl 水分两次浸洗 PVDF 膜, 合并水解液与浸洗液, 离心冷冻干燥, 放置 -20℃待衍生分析中性糖。

(6) 单糖 AMAC 衍生

向已冻干的单糖样品中加 5μl 50mmol/L AMAC(溶于 HAc/DMSO, 3/17)及 5μl 1.0mol/L $NaBH_3CN$(溶于 THF), 振荡, 稍离心。唾液酸的衍生反应条件是在 55℃水浴保温反应 4h, 其他单糖的衍生反应条件是 45℃水浴保温反应 5h。反应完毕后, 离心冷冻干燥, 加适当体积(50～400μl)的 50% DMSO/H_2O 溶解, 用毛细管电泳分析。

(7) 寡聚葡萄糖混合物的制备及寡糖 ANTS 衍生

置 0.2g 葡聚糖(Dextran)于一安瓿瓶底部, 加 1.5ml 0.1mol/L HCl, 充高纯氮 1min, 封管, 在 100℃恒温水解 4h。离心冷冻干燥水解产物至干。加 1ml Milli-Q 水, 振荡, 离心冷冻干燥至干以除去酸, 得到寡聚葡萄糖混合物。取其 50mg 于 0.5ml Eppendorf 管, 加 150μl 0.2mol/L ANTS(溶于 HAc/H_2O, 3/17), 200μl 1.0mol/L $NaBH_3CN$(溶于 DMSO), 振荡, 密封, 置 40℃水浴反应 15h。取 5μl 用 Milli-Q 水稀释至 400μl, 用毛细管电泳分析。

(8) 毛细管电泳分析单糖 AMAC 衍生物

单糖 AMAC 衍生物分离在 ABI 270A 型毛细管电泳仪上进行。使用未涂层熔融石英毛细管, 总长度及有效长度(进样端到检测器处长度)见图注, 内径 50μm。电泳缓冲液为 100mmol/L $Na_2B_4O_7$, pH10.5(用 NaOH 调 pH 值), 或 300mmol/L H_3BO_3, pH10.5(用 NaOH 调 pH 值)。电泳电压见图注, 毛细管恒温在(30±0.2)℃。检测波长为 264 或 260nm。电泳图和数据收集由惠普 HP3394A 积分仪完成。每次电泳前分别用 0.1mol/L NaOH 和 Milli-Q 水冲洗毛细管 2min, 再用电泳缓冲液冲洗平衡 5min。

(9) 毛细管电泳分析寡糖 ANTS 衍生物

分析寡糖 ANTS 衍生物时, 电泳缓冲液为 50mmol/L 磷酸缓冲液, pH2.5, 电泳电压 -20kV 或使用 100mmol/L, pH9.3 的硼砂缓冲液为电泳缓冲液, 电泳电压 15kV。检测波长 214nm。其余条件同方法 7。

四、结果和讨论

图 11.11 是 9 种还原单糖 AMAC 衍生物在 100mm/L $Na_2B_2O_7$, pH10.5 缓冲液中毛

细管电泳分离图谱。各峰归属通过混合单糖 AMAC 衍生物与每一单糖 AMAC 衍生物共

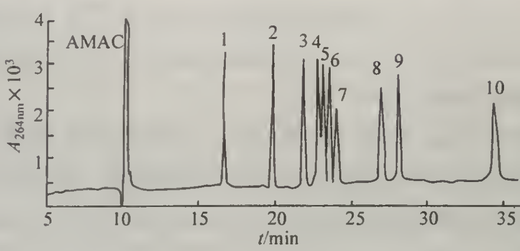


图 11.11　毛细管电泳分离 9 种单糖 AMAC 衍生物

条件：熔融石英毛细管，内径 50μm，总长 62cm，有效长度 40cm；电泳缓冲液，100mmol/L $Na_2B_4O_7$-NaOH，pH10.5；电泳电压为 14kV；检测波长 264nm；单糖 AMAC 衍生物浓度均为 20μmol/L，真空进样 2s(约 6 nl)。各峰归属：1. N-乙酰氨基半乳糖；2. 木糖；3. 核糖；4. N-乙酰氨基葡萄糖；5. 葡萄糖；6. 甘露糖；7. 阿拉伯糖；8. 岩藻糖；9. 半乳糖；10. 肉桂酸(内标)

进样而确定。不同单糖 AMAC 衍生物在硼酸盐缓冲液中电迁移率的差异原因，也是毛细管电泳分离单糖 AMAC 衍生物的基础，主要是由于不同化学结构的单糖与硼酸根 $B(OH)_4^-$ 形成的络合物稳定性不同。越稳定的络合物迁移率越大，而稳定性与单糖部分邻位强基的取向、邻位强基的个数有关[16]。单糖浓度在 5～25μmol/L 内，各单糖 AMAC 衍生物峰面积与迁移时间之商关于单糖浓度成良好的线性相关，相关系数均大于 0.992，表明该方法可定量分析单糖。图 11.12A 和图

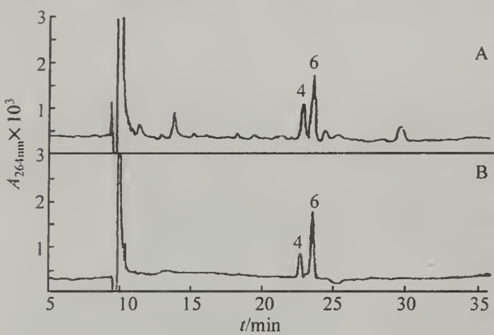


图 11.12　毛细管电泳分析糖蛋白中单糖组成

A. 火鸡卵白蛋白(A5015)；B. 核糖核酸酶 B。毛细管电泳条件和各峰归属同图 11.11

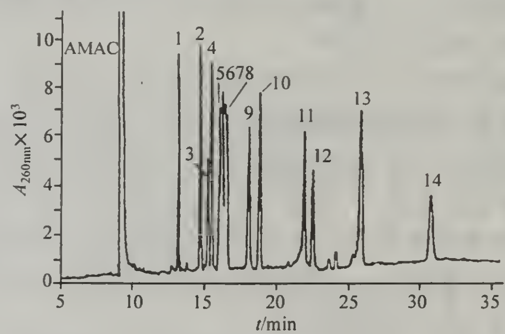


图 11.13　毛细管电泳分析酸性单糖 AMAC 衍生物

条件：熔融石英毛细管，内径 50μm，总长 72cm，有效长度 50cm；电泳缓冲液，300mmol/L H_3BO_3-NaOH，pH10.5；电泳电压为 20kV；检测波长 260nm；唾液酸AMAC衍生物浓度为 100μmol/L，其余单糖 AMAC 衍生物浓度均为 25μmol/L，真空进样 2s(约 7nl)。各峰归属：1. N-乙酰氨基半乳糖；2. 木糖；3. 唾液酸；4. 核糖；5. N-乙酰氨基葡萄糖；6. 葡萄糖；7. 甘露糖；8. 阿拉伯糖；9. 岩藻糖；10. 半乳糖；11. 肉桂酸(内标)；12. 唾液酸；13. 葡萄糖醛酸；14. 半乳糖醛酸

11.12B 分别是分析火鸡卵白蛋白(A5015)和核糖核酸酶 B 糖成分的电泳图，显示这两种糖蛋白糖主要成分为 GlcNAc 和 Man,定量分析它们的摩尔比分别为 2:4.3 和 2:6.6 。

图 11.13 是在 300mmol/L H_3BO_3,pH10.5(用 NaOH 调 pH)电泳缓冲液中分离中性单糖和酸性单糖 AMAC 衍生物的电泳图。需要说明的是,由于该电泳条件是用来分析酸性单糖(主要是分析唾液酸)[33],所以,*N*-乙酰氨基葡萄糖、葡萄糖、甘露糖和阿拉伯糖这 4 种糖的 AMAC 衍生物未被很好分离并不重要。另外,唾液酸 AMAC 衍生物,如 5-*N*-乙酰氨基唾液酸 Neu5Ac-AMAC(图 11.13 中,峰 12)在室温下不稳定,易自动转化成稳定的脱羧基唾液酸 AMAC 衍生物,即脱羧基 5-*N*-乙酰氨基唾液酸 AMAC 衍生物(图 11.13 中,峰 3)。而在 −20℃ 暗处保存 12h,基本观测不到唾液酸 AMAC 衍生物的脱羧反应。在 10～120μmol/L 浓度范围,唾液酸 AMAC 衍生物的峰面积与迁移时间之商与唾液酸浓度成良好的线性关系,线性相关系数为 0.9978($n=8$)。表明在所给的条件下,可以准确定量唾液酸。图 11.14 A,B 和 C 分别是分析重组人促红细胞生成素 rhu-EPO,人尿胰蛋白酶抑制剂(hu-UTI)和牛酸性糖蛋白(α_1-AGP)中唾液酸的毛细管电泳图。图 11.14 C 显示牛 α_1-AGP 中除存在 Neu5Ac 外,还存在另一酸性糖。它的AMAC衍生物(带*号峰)也能发生脱羧反应。用质谱鉴定这一酸性糖为 5-*N*-乙醇酰氨基唾液酸(Neu5Gc)[33]。定量分析表明,每 mol rhu-EPO 和 hu-UTI 分别含 17.59 和 2.55mol Neu5Ac,而每 mol 牛 α_1-AGP含 5.5mol Neu5Ac 和 5.04mol Neu5Gc。

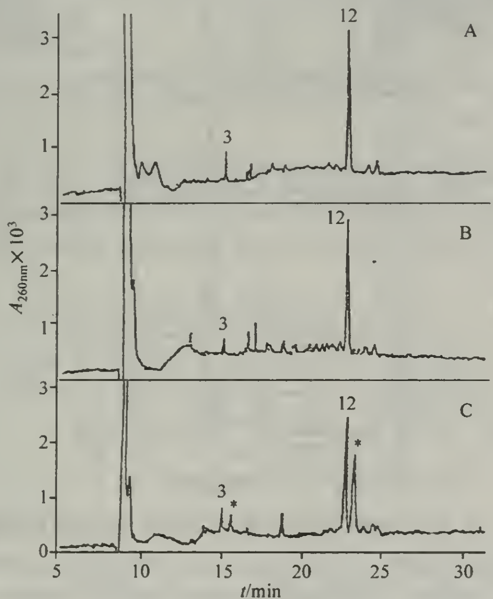


图 11.14　毛细管电泳分析糖蛋白中唾液酸

酸水解释放糖蛋白中的唾液酸经 AMAC 衍生化,衍生物溶于 100μl 50% DMSO/H_2O。A.0.45nmol rhu-EPO,真空进样 2s; B.2.14nmol hu-UTI,真空进样 2s; C.1.42nmol 牛 α_1-AGP,真空进样 1s。其余条件和各峰归属同图 11.13

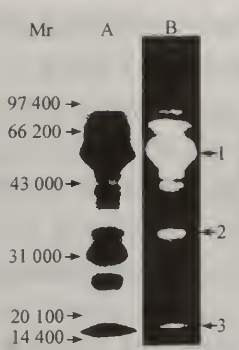


图 11.15 荧光检测 SDS-PAGE 分离并电转移于 PVDF 膜上蛇毒糖蛋白

A. PVDF 膜上蛋白考马斯亮蓝染色；B. PVDF 膜上糖蛋白荧光染色

江浙蝮蛇蛇毒硫酸铵反透析沉淀总蛋白等分为二，分别上样于 SDS-PAGE 胶两样品道，电泳分离后，电转移蛋白质至 PVDF 膜。从两样品道中间剪开 PVDF 膜，一半 PVDF 膜用糖蛋白荧光染色法染色[44]，另一半 PVDF 膜用考马斯亮蓝染色。荧光染色结果（图 11.15）表明有若干蛋白条带被染色，说明蝮蛇中存在较多的糖蛋白。用 CE 对图中标示出的 3 条带的单糖成分进行分析，证实了它们是糖蛋白。图 11.16A 是单糖AMAC 衍生物标准的毛细管电泳分离图谱。图 11.16B 是全程空白对照实验，可知 2mol/L TFA 水解液存在木糖和葡糖糖。图 11.16 C，D 和 E 分别是分子质量约 55，32，14kDa 糖蛋白的中性糖分析结果，表明它们都含有甘露糖、半乳糖和岩藻糖。由于我们没有对在水解产物进行重乙酰化，因此没有检测到 *N*－乙酰氨基葡萄糖和 *N*－乙酰氨基半乳糖。唾液酸分析结果（图谱未给出）表明没有检测到唾液酸。

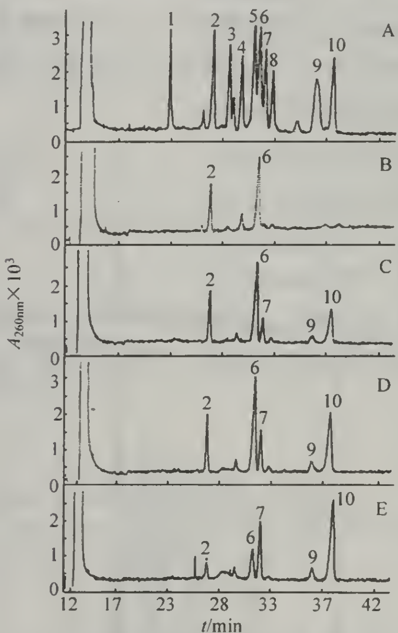


图 11.16 毛细管电泳分析 PVDF 膜上糖蛋白的单糖

条件：熔融石英毛细管，内径 50μm，总长 72cm，有效长度 50cm；电泳缓冲液，100mmol/L $Na_2B_4O_7$-NaOH，pH10.5；电泳电压为 17kV；检测波长 260nm；

A. 10 种单糖 AMAC 衍生物标准毛细管电泳图。各峰归属：1. *N*－乙酰氨基半乳糖；2. 木糖；3. 唾液酸；4. 核糖；5. *N*－乙酰氨基葡萄糖；6. 葡萄糖；7. 甘露糖；8. 阿拉伯糖；9. 岩藻糖；10. 半乳糖；B. 全程空白对照实验；C. No.3 糖蛋白的单糖；D. No.2 糖蛋白的单糖；E. No.1 糖蛋白的单糖

图 11.17A 为在 50mmol/L，pH2.5 的磷酸缓冲液中葡聚糖酸水解产物，即不同聚合度寡聚葡糖（图 11.18）混合物的 ANTS 衍生物电泳图。聚合度从 1～21 的寡聚葡糖 ANTS 衍生物得到检测和近基线分离。在 pH2.5 的缓冲溶液条件下，毛细管内表面硅烷

醇的氢解离性很弱,电渗流很弱,而 ANTS 寡糖衍生物带近 3 个负电荷。所以这时电迁移速度远大于电渗流速度,而且两者的方向相反。因此,需在阳极检测衍生物,也即施加负极性电压,使阳极在检测器一侧。出峰顺序按聚合度 1,2,3,…21 排列。这被毛细管电泳-离子阱质谱联用分析糖 ANTS 衍生物技术证实[45]。图 11.17B 是在 100mmol/L $Na_2B_4O_7$,pH9.3 缓冲液中寡聚葡糖 ANTS 衍生物的电泳图。在碱性缓冲液中,电渗流速度很快,需要在阴极检测 ANTS 衍生物,而且出峰顺序正好与图 11.17B 的相反。很明显,这时聚合度越大的衍生物之间的分离度越小,这是因为聚合度越大的衍生物之间的电迁移速度的差异比电渗流速度越来越小,在衍生物到达检测器时尚未得到有效分离。寡糖 ANTS 衍生物在 pH2.5 磷酸缓冲液中的电迁移率差异主要决定于寡糖分子的有效表面的差异。因此,可以通过比较结构相似的寡糖 ANTS 衍生物之间的电迁移率的差异得到寡糖相对大小的信息[32]。

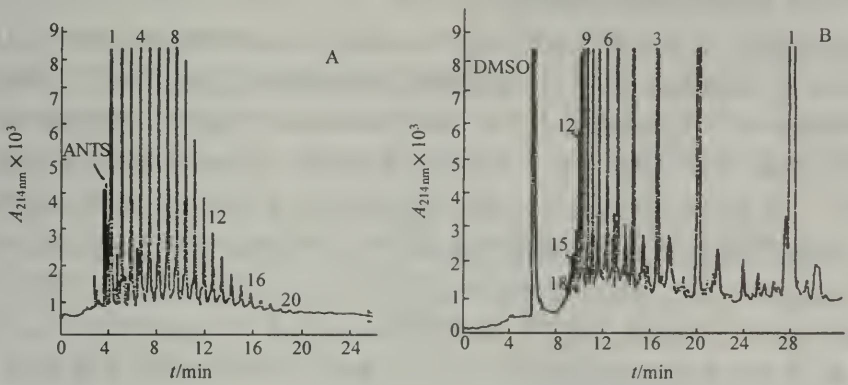


图 11.17　毛细管电泳分离葡聚寡糖 ANTS 衍生物

条件:熔融石英毛细管,内径 50μm,总长 60cm,有效长度 40cm;检测波长 214nm。A. 电泳缓冲液 50mmol/L $NaH_2PO_3-H_3PO_3$,pH2.5;电泳电压为 -20kV;B. 电泳缓冲液 100mmol/L $Na_2B_4O_7$,pH9.3;电泳电压为 15kV

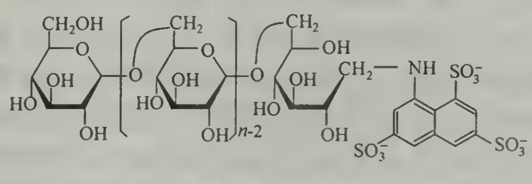


图 11.18　葡聚寡糖 ANTS 衍生物结构

(n = 1, 2, 3, …)

第三节　毛细管电泳分析糖蛋白微不均一性

一、简　　介

糖蛋白一个特点就是它趋于“微不均一性”。某一种糖蛋白实际上是一类具有相同的多肽链,而共价连接于肽链上的寡糖的结构不同(包括单糖成分和连接方式)的蛋白群体。不同结构的寡糖链,即不同糖型使糖蛋白表现出微不均一性。

毛细管电泳(CE)的高分离能力使分离糖蛋白的糖型成为可能。它为回答诸如糖蛋白糖基部分的生物功能;与自身的非糖基化蛋白相比较,不同的糖型是否具有不同的生物或生理活性;某些糖蛋白的糖型的变化是否与某一疾病有相关性等问题提供了一种有效的技术支持。同时也为生物医药工业重组糖蛋白的质量控制提供了一种快速分析手段。

目前,已有很多文献报道 CE 分析糖蛋白微不均一性。几种分离模式,如毛细管区带电泳(CZE)、毛细管等电聚焦(CIEF)和微团电动毛细管色谱(MEKC)[46]等都被应用来研究糖蛋白微不均一性 。

最早报道 CE 分析糖蛋白微不均一性的是 Kilar 和 Hjerten[47]。他们用 CZE 方式分离人血清转铁蛋白的糖型,得到 5 个组分峰,分别代表含 2,3,4,5 和 6 个唾液酸残基的转铁蛋白。在用唾液酸水解酶处理转铁蛋白不同时间后,用 CE 可以观测这 5 个糖型峰高发生规律的变化,表明唾液酸逐渐被水解掉。处理 20h 后,基本只有一个新的峰存在。用 CIEF 鉴定该峰是去唾液酸化的转铁蛋白。

医药用的重组人促红细胞生成素(rhEPO)的糖部分结构相当复杂,有 4 位点被糖基化修饰,既有 *N*-连接糖链,也有 *O*-连接糖链,其唾液酸的含量也很高[33]。利用 CE 研究 rhEPO 的微不均一性的报道较多。在 pH4.0,100mmol/L 醋酸钠-醋酸缓冲液中 CE 分析 rhEPO 仅得一个峰,而在 pH4.0,100mmol/L 醋酸钠-磷酸缓冲液中电泳分离得到 5 个组分峰[48]。其原因可能是磷酸与毛细管内壁硅醇形成复合物,使毛细管对蛋白质的吸附降低。表明缓冲液组成成分在糖蛋白糖型的分析中所起的重要作用。在 10mmol/L Tricine, 10mmol/L NaCl,pH6.2 缓冲液中,添加 2.5mmol/L 1,4-二氨基丁烷后,使原来仅一个峰的 rhEPO 分离成 4 个峰[49]。再添加 7mol/L 尿素于电泳缓冲液中,rhEPO 被分离成 6 个峰,并且分离度明显提高很多。1, 4-二氨基丁烷作用是降低电渗流,尿素可能抑制或降低了毛细管内壁对 rhEPO 的吸附。另外,需注意,在 7mol/L 尿素存在下,蛋白质处于变性状态。而应用 CIEF 分析 rhEPO 糖型[50],在 7mol/L 尿素和两性电解质(pH 3~10两性电解质与 pH 2.5~5 的两性电解质以体积比 1:2 混合)中等电聚焦,得到最佳分离。在表观等电点 3.78~4.69 范围内,可观测到 7 个糖型。

用 CE 分析鸡卵白蛋白的糖型的报道也较多。在 100mmol/L 硼酸盐,pH8.5 的缓冲液中用 CZE 方式分离鸡卵白蛋白,仅观测到几个峰。而当缓冲液中存在 1mmol/L 1,4-二氨基丁烷时,可观察到 20 多个大小不同的峰[51]。用键合交联聚丙烯酰胺涂层的毛细管,在 0.4mol/L 硼酸,2mol/L Tris,2mmol/L EDTA,pH7.8 缓冲液中分离鸡卵白蛋白,也可得到 20 多个峰[52]。毛细管内壁中性涂层同样是为了降低电渗流并抑制蛋白质与毛细管内壁相互作用,以达到高分辨的目的。总之,可从如何降低电渗流,如何降低毛细管内壁对糖蛋白的吸附,以及从缓冲液的 pH 值、组成和浓度等方面考虑和尝试,选择糖型分析的条件。修饰毛细管内表面以降低电渗流和降低表面对蛋白质的吸附的方法分为动态涂层和静态涂层两类。动态涂层是在电泳缓冲液中加多聚物、离子或非离子有机去垢剂、二胺或多胺等。这些添加剂基于吸附原理涂附于毛细管内壁,因此涂层方便并易再生,而且电泳分离重现性高。静态涂层是将毛细管内壁经共价反应“永久”键合涂敷一层中性亲水多聚物或带电荷亲水多聚物或疏水多聚物。静态涂层通常稳定(但在强酸或强碱中稳定性下降),对抑制和降低毛细管电渗流和表面对蛋白质的吸附更有效;但制备耗时。若欲了解有关动态或静态涂层详细信息,可参阅相关文献综述[53~55]。最近,动态

涂层(电泳缓冲液 100mmol/L 硼酸盐中加 3mmol/L 二氨基丁烷)和静态涂层(碳氟化合物涂层)毛细管电泳都被运用来直接分析人血清中转铁蛋白唾液酸糖型[56, 57],并观测到嗜酒和非嗜酒病人血清中转铁蛋白唾液酸糖型的差异。两种方法得到可比较的糖型分离结果。

需要指出的是,由于糖蛋白结构极其复杂,不同糖蛋白间的差异,包括多肽链一级结构和空间结构(包括二硫键)的差异、糖基化位点的多少、糖含量的多少、寡糖链结构的差异、有无唾液酸和唾液酸的含量多少等,目前很难有 1 或 2 种固定的分离模式或分离条件对所有或大多数糖蛋白都得到满意的糖型分离结果。因而,还必需根据待分析糖蛋白的特性,寻找最佳的、可重现的、且易操作的分离条件。例如,无论是用涂层或是未涂层的毛细管,在硼酸缓冲液中,即使有二氨基丁烷存在于缓冲液,也不能将重组人组织纤溶酶原激活物(rh-tPA)糖型分离[58],表明唾液酸化程度高的糖蛋白不能有效地与硼酸根形成复合物,并且糖型间的电迁移率差异甚小,因而难以得到分离。唾液酸含量不同的糖型由于其等电点不同,可用 CIEF 将它们分离。如用 pH3~10 和 pH5~8 两性电解质混合物(体积比 1:1)为等电聚焦电泳液(其中还含 4mol/L 尿素),可将 rh-tPA 分离成 10 个峰[59]。使用尿素的目的是增加蛋白质的溶解性,因为蛋白质在其等电点处可能沉淀。而对于只有一个位点被糖基化并且糖链是高甘露糖型的重组人骨形态形成蛋白 2(rhBMP-2),使用简单的电泳缓冲液 0.1mol/L,pH2.5 磷酸缓冲液及中性涂层毛细管,就可将其 9 个糖型分离[60]。

下面介绍我们用 CZE 比较分析 3 种卵白蛋白糖型分布的方法和结果[61]。

二、材　　料

两种火鸡卵白蛋白(Sigma 产品 A5015,A7269)和鸡卵白蛋白(A2512),1,4 -二氨基丁烷,异亚丙基丙酮(mesityloxide)。

三、方　　法

糖型分析在 ABI 270A 型毛细管电泳仪上进行。使用总长 72cm,有效分离长度 50cm,内径 50μm 熔融石英毛细管(ABI 产品)。检测波长为 200nm。毛细管温度控制在(30±0.2)℃。进样前分别用 0.1mol/L NaOH 和水冲洗 2min,再用电泳缓冲液冲洗平衡 4min。使用的电泳缓冲液为:(1)50mmol/L NaH_2PO_4-H_3PO_4, pH2.5;(2)100mmol/L H_3BO_3-NaOH,pH8.6;(3)100mmol/L $Na_2B_4O_7$,40~100mmol/L KH_2PO_4,pH8.2~9.0;(4)100mmol/L H_3BO_3-NaOH, pH8.6, 1.8mmol/L 1,4 -二氨基丁烷。电泳电压随缓冲液而调整。用电中性物质异亚丙基丙酮测定毛细管电渗流。

四、结果和讨论

商品化的两种火鸡卵白蛋白(A5015,A7269)和鸡卵白蛋白(A2512)经反相 HPLC 进一步纯化。在 50mmol/L 磷酸,pH2.5 缓冲溶液中分析这 3 种卵白蛋白,仅微弱显示它们均存在微不均一性。在 100mmol/L H_3BO_3-NaOH,pH8.6 缓冲液中分离 A5015 效果有所

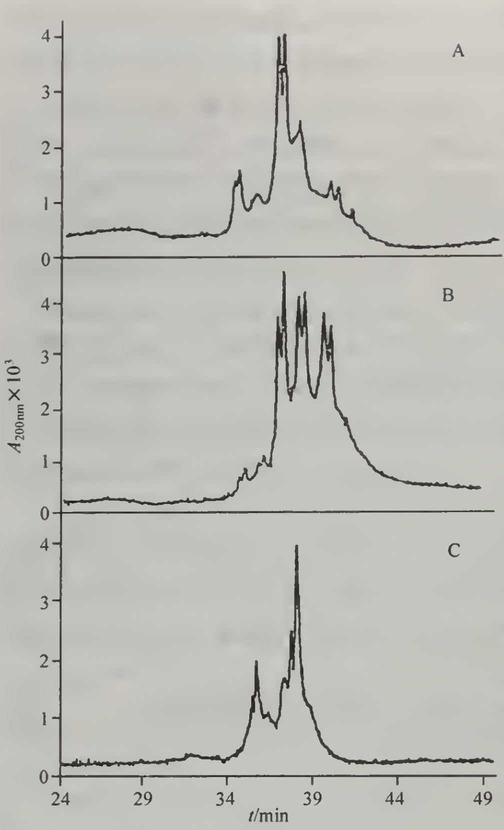


图 11.19　高离子强度硼酸盐缓冲液中分离卵白蛋白糖型

条件：熔融石英毛细管，内径 50μm，总长 72cm，有效长度 50cm；检测波长 200nm。A. 火鸡卵白蛋白 A5015。电泳缓冲液，100mmol/L $Na_2B_4O_7$，100mmol/L KH_2PO_4，pH8.2；电泳电压为 8kV。B. 火鸡卵白蛋白 A7269。电泳缓冲液，100mmol/L $Na_2B_4O_7$，40mmol/L KH_2PO_4，pH9.0；电泳电压为 8kV。C. 鸡卵白蛋白 A2512。电泳缓冲液，100mmol/L $Na_2B_4O_7$，60mmol/L KH_2PO_3，pH8.8；电泳电压为 10kV

增加，可以观察到 5 个峰。然而由于在碱性缓冲液中毛细管的电渗流变得较大，蛋白在电渗流的负载下较快地通过检测器而没有得到充分的电泳分离。基于电渗流的性质，通常有 3 种方法消除或抑制电渗流：一是使用中性涂层毛细管；二是使用高离子强度的缓冲液；第三种方法是使用化学添加剂，如有机溶剂或有机阳离子试剂于电泳缓冲液中。

图 11.19 是 3 种卵白蛋白在较高离子强度硼砂缓冲液中的毛细管电泳图。这时电渗流降低，糖型分离度进一步提高。A5015 和 A7269 已被分离成 5 组峰，A2512 被分离成两组峰。但是，继续提高硼砂或磷酸二氢钾的浓度，并不能继续提高糖型的分离度。另外，如继续增加电压（>150V/cm），会降低那些相邻的峰的分辨率。这可归结于高离子强度和高电压，导致大量的焦尔热产生而引起峰严重变宽甚至坍陷。因此，当使用高离子强度缓冲液时，施加在毛细管两端的电压被限制在某一值以下，最终阻碍了分离度进一步提高。分别在含不同浓度的 1,4-二氨基丁烷（0～2.0mmol/L）的 100mmol/L H_3BO_3-NaOH，pH8.6 缓冲液中，进行分离 A5015 糖型的一系列实验，发现当 1,4-二氨基丁烷浓度为 1.8mmol/L 时，A5015 糖型得到最佳分离。图 11.20A，B 分别是在最适分离条件（100mmol/L H_3BO_3-NaOH，pH8.6，1.8mmol/L 1,4-二氨基丁烷，20kV）下，A5015 和 A7269 的糖型分离图，多于 13 个峰在 35min 内可被较好分离。对 A7269 连续 6 次电泳得到的各糖型峰的相对迁移时间（相对于 EOF）变异系数均在 2.7% 以下，表明分离的可重复性较高。有趣的是，A5015 和 A7269 的糖型花样极其相似，每一对相对应的峰的相对迁移时间几乎相同。将 A5015 和 A7269 按 1∶1 混合，电泳发现它们的每一对相对应的峰完全重叠（图 11.20C）。这表明 A5015 和 A7269 具有完全相同的糖型花样。它们之间的不同主要在于峰 8，9，12，13 的相对含量不同。与 A7269 相比，A5015 中糖型 8，9，12，13 的相对含量明显降低。图 11.20D 为在上述分离火鸡卵白蛋白最适条件下，分离鸡卵白蛋白 A2512 糖型电泳图。比较火鸡卵白蛋白和鸡卵白蛋白糖型电泳图，可预知其寡糖链结构不同。

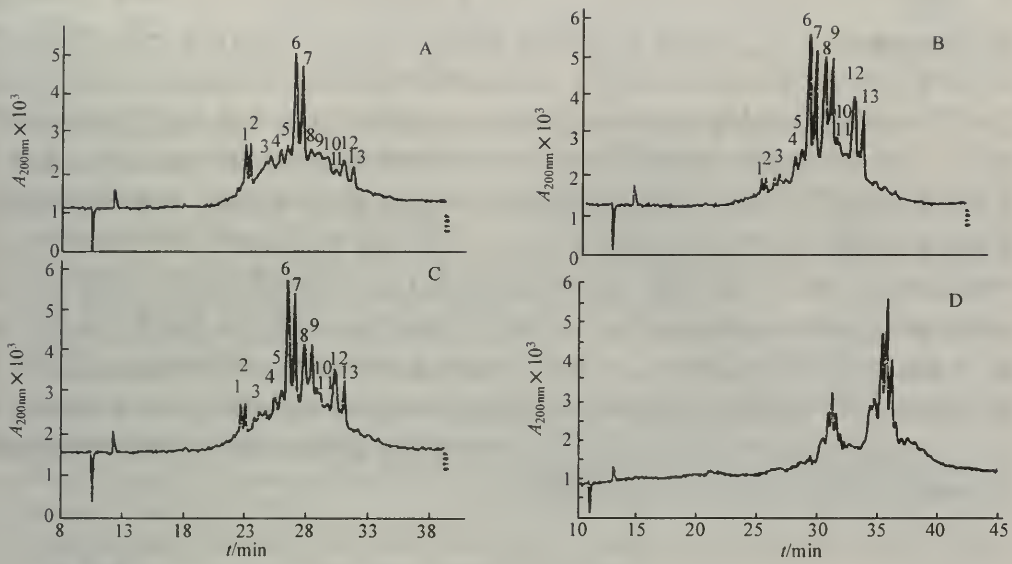


图 11.20　含 1,4-二氨基丁烷硼酸盐缓冲液中分离卵白蛋白糖型

条件:熔融石英毛细管,内径 50μm,总长 72cm,有效长度 50cm;检测波长 200nm。电泳缓冲液,100mmol/L H_3BO_3-NaOH,pH8.6,1.8mmol/L 1,4-二氨基丁烷。电泳电压为 20kV。蛋白浓度,1mg/L,真空进样 1s。A. 火鸡卵白蛋白 A5015;B. 火鸡卵白蛋白 A7269;C. 火鸡卵白蛋白 A5015 和 A7269 混合物(1:1);D. 鸡卵白蛋白 A2512

第四节　结语和展望

目前,由于质谱技术的迅速发展,蛋白质磷酸化分析多用质谱完成[62~64]。但在 pSer 或 pThr 残基相邻有多个 Ser 或 Thr 残基存在情况下,用质谱确定是丝氨酸或苏氨酸磷酸化通常比较困难。结合已知的蛋白激酶底物识别序列(recognition motifs)推测磷酸化位点,有时得到错误结论。因为至今有许多蛋白激酶还没被发现,其底物识别序列更是未知的。而用 CE 分析磷酸氨基酸方法容易区分磷酸丝氨酸和磷酸苏氨酸。当使用荧光[65]或激光诱导荧光检测器,CE 分析磷酸氨基酸方法的灵敏度会得到较大提高。对于 CE 分析糖分子,已有若干糖衍生法和衍生试剂相继被开发出来,而目前最常用的是还原胺化法。最近,又有一种糖衍生法被报道[66]。衍生第一步是还原糖与甲基胺还原胺化反应生成 *N*-甲基糖胺(*N*-methylglycamine);第二步中,*N*-甲基糖胺与 4-nitro-2,1,3-benzoxadiazole 7-fluoride(NBD-F)反应生成 NBD 标记的 *N*-甲基糖胺(NBD-tagged *N*-methylglycamine)。衍生反应在 pH4.5, 40℃进行 30min 定量完成。另一方面,各种检测方法得到运用,灵敏度以激光诱导荧光(LIF)检测最高,但目前价格贵。另外,由于激光发生器产生的激光波长有限,应用受到一定的限制。紫外检测法较常用,检测波长范围宽,是灵敏度与价格的折衷选择。CE 蛋白糖基化修饰分析涉及糖蛋白单糖组成分析、寡糖分析、寡糖序列分析[40]、糖肽分析、糖蛋白糖型分析等方面。

过去 12 年里,毛细管电泳技术得到快速发展,在生命科学中发挥了巨大的作用。人类基因组计划进展迅猛,几近完成,可以说得益于高通量毛细管电泳 DNA 测序技术的运

用。而近 6～7 年也是全微分析系统(micro-total analysis systems, μ-TAS)或称“芯片实验室”(“lab-on-a chip”)[67]研究真正兴起和发展时期。这种 μ-TAS 将完成分离和检测的所有部件整合在一个微小化平台装置内。目前毛细管电泳的微小化是最成功的微芯片应用之一[68]。在微芯片长方形管道中的毛细管电泳分离情况与在开管毛细管电泳是相同的[69]。几种毛细管电泳分离模式如 CZE, CGE(毛细管凝胶电泳)和 MEKC 都已在微芯片电泳中得到应用。最近,文献[70]报道微芯片电泳快速分析氨基糖。氨基糖(氨基葡萄糖和氨基半乳糖)及它们的还原形式经 NBD-F 衍生,衍生物在微制作芯片 33mm 长的分离管道内电泳分离,用 LIF 检测。这 4 种衍生物在 1min 内得到很好分离。低至 0.5fmol 的氨基糖可被检测到(信噪比 S/N=3)。微芯片毛细管电泳可望发展成高度自动化、大规模、高通量、快速的分离分析方法,除可配以光或电化学或拉曼光谱等检测器外,还可以质谱为检测器[71,72],增强其分离分析能力,获得更多的被分析物质的信息。可以相信,这种整合方式可能更有助于后基因组时代,进行大规模高通量蛋白质翻译后修饰,特别是磷酸化和糖基化修饰的分析鉴定以及临床诊断。

(车发云)

参考文献

1 Hubbard MJ, Cohen P. On Target with a new mechanism for the regulation of protein phosphorylation. Trends Biochem Sci, 1993, 18: 172～177

2 Hunter T. Protein kinase classification. Methods Enzymol, 1991, 200: 3～37

3 Cohen P. Signal integration at the level of protein kinases, protein phosphatases and their substrates. Trends Biochem Sci, 1992, 17: 408～413

4 Fadden P, Haystead TA. Quantitative and selective fluorophore labeling of phosphoserine on peptides and proteins: characterization at the attomole level by capillary electrophoresis and laserinduced fluorescence. Anal Biochem, 1995, 225: 81～88

5 Cooper JA, Stefton BM, Hunter T. Detection and quantification of phosphotyrosine in proteins. Methods Enzymol, 1983, 99: 387～402

6 Hanks SK, Quinn AM, Hunter T. The protein kinase family: conserved features and deduced phylogeny of the catalytic domains. Science, 1988, 241: 42～52

7 Meyer HE, Hoffmann-Posorske E, Donella-Deana A *et al* Sequence analysis of phosphotyrosine-containing peptides. Methods Enzymol, 1991, 201: 206～224

8 车发云, 邵晓霞, 徐来根等. 毛细管电泳非放射分析蛋白质和多肽磷酸化. 生物化学与生物物理学报, 1997, 29: 363～370

9 万谦, 陈长征, 屠红旻等. 毛细管电泳测定蛋白激酶 A 活力的新方法. 生物化学与生物物理学报, 1995, 27: 579～584

10 Bylund DB, Huang TS. Decomposition of phosphoserine and phosphothreonine during acid hydrolysis. Anal Biochem, 1976, 73: 477～485

11 Kornfeld R, Kornfeld S. Assembly of asparagine-linked oligosaccharides. Ann Rev Biochem, 1985, 54: 631～664

12 O'Neill RA. Enzymatic release of oligosaccharides from glycoproteins for chromatographic and electrophoretic analysis. J Chromatogr A, 1996, 720: 201～215

13 El Rassi Z. Recent developments in capillary electrophoresis and capillary electrochromatography of carbohydrate species. Electrophoresis, 1999, 20: 3134～3144

14 Honda S, Iwase S, Makino A. Simultaneous determination of reducing monosaccharides by capillary zone electrophore-

sis as the borate complexes of N-2-pyridylglycamines. Anal Biochem, 1989, 176: 72～77
15 常理文，腰锐峰，陈义. 单糖的高效毛细管区带电泳定量分析. 分析化学，1994,22: 1～5
16 车发云，刘正宇，王克夷等. 毛细管电泳分析常见单糖的 2-氨基吖啶酮衍生物. 生物化学与生物物理学报，1998,30: 251～256
17 Liu JP, Shirota O, Wiesler D *et al*. Ultrasensitive fluorometric detection of carbohydrates as derivatives in mixtures separated by capillary electrophoresis. Proc Natl Acad Sci USA, 1991, 88: 2302～2306
18 Chen FT, Dobashi TS, Evangelista RA. Quantitative analysis of sugar constitutents of glycoproteins by capillary electrophoresis. Glycobiology, 1998, 8: 1045～1052
19 Hoffstetter-Kuhn S, Paulus A, Gassmann E *et al*. Influence of borate complexation on the electrophoretic behavior of carbohydrates in capillary electrophoresis. Anal Chem, 1991,63: 1541～1547
20 Dong X, Xu X, Han F *et al*. Determination of sialic acids in the serum of cancer patients by capillary electrophoresis. Electrophoresis, 2001,22: 2231～ 2235
21 Klockow A, Paulus A, Figueiredo A *et al*. Determination of carbohydrates in fruit juices by capillary electrophoresis and high-performance liquid chromatography. J Chromatogr A, 1994, 680: 187～200
22 Soga T, Heiger DN. Simultaneous determination of monosaccharides in glycoproteins by capillary electrophoresis. Anal Biochem, 1998,261: 73～78
23 Colón LA, Dadoo R, Zare RN. Determination of carbohydrates by capillary zone electrophoresis with amperometric detection at a copper microelectrode. Anal Chem, 1993, 65: 476～481
24 Honda S, Suzuki S, Nose A *et al*. Capillary zone electrophoresis of reducing mono-and oligo-saccharides as the borate complexes of their 3-methl-1-phenyl-2-pyrazolin -5-one derivatives. Carbohydr Res, 1991, 215: 193～198
25 Mechref Y, El Rassi Z. Capillary zone electrophoresis of derivatized acidic monosaccharides. Electrophoresis, 1994,15: 627～634
26 Suzuki S, Kakehi K, Honda S. Two-dimensional mapping of N-glycosidically linked asialo-oligosaccharides from glycoproteins as reductively pyridylaminated derivatives using dual separation modes of high-performance capillary electrophoresis. Anal Biochem, 1992,205: 227～236
27 Grill E, Huber C, Oefner PJ *et al*. Capillary zone electrophoresis of paminobenzoic acid derivatives of aldoses, ketoses and uronic acids. Electrophoresis, 1993,14: 1004～1010
28 Schwaiger H, Oefner PJ, Huber C *et al*. Capillary zone electrophoresis and micellar electrokinetic chromatography of 4-aminobenzonitrile carbohydrate derivatives. Electrophoresis, 1994,15: 941～952
29 Stefansson M, Novotny M. Separation of complex oligosaccharide mixtures by capillary electrophoresis in the open-tubular format. Anal Chem, 1994,66: 1134～1140
30 Klockow A, Amado R, Widmer HM *et al*. Separation of 8-aminonaphthalene-1, 3, 6-trisulfonic acid-labelled neutral and sialylated N-linked complex oligosaccharides by capillary electrophoresis. J Chromator A, 1995,716: 241～247
31 车发云，刘正宇，王克夷等. 毛细管电泳分离寡糖衍生物及其电泳行为研究. 生物化学与生物物理学报，1998, 30: 495～500
32 车发云，王克夷，夏其昌. 寡糖 8-氨基萘基-1, 3, 6-三磺酸衍生物在毛细管区带电泳中迁移时间相互关系. 中国生物化学与分子生物学报，1999,15: 126～129
33 Che FY, Shao XX, Wang KY *et al*. Characterization of derivatization of sialic acid with 2-aminoacridone and determination of sialic acid content in glycoproteins by capillary electrophoresis and high performance liquid chromatography-ion trap mass spectrometry. Electrophoresis, 1999,20: 2930～2937
34 Liu J, Shirota O, Novotny M. Sensitive, laser-assisted determination of complex oligosaccharide mixtures separated by capillary gel electrophoresis at high resolution. Anal Chem, 1992,64: 973～975
35 Guttman A. Analysis of monosaccharide composition by capillary electrophoresis. J Chromatogr A, 1997,763: 271～277
36 Zhao JY, Diedrich P, Zhang Y *et al*. Separation of aminated monosaccharides by capillary zone electrophoresis with laser-induced fluorescence detection. J Chromatogr B Biomed Appl, 1994,657: 307～313

37 Starr CM, Masada RI, Hague C *et al*. Fluorophore-assisted carbohydrate electrophoresis in the separation, analysis, and sequencing of carbohydrates. J Chromatogr A, 1996,720: 295～321

38 Paulus A, Klockow A. Detection of carbohydrates in capillary electrophoresis. J Chromatogr A, 1996,720: 353～376

39 Prime S, Dearnley J, Ventom AM *et al*. Oligosaccharide sequencing based on exo and endoglycosidase digestion and liquid chromatographic analysis of the products. J Chromatogr A, 1996,720: 263～274

40 Guttman A, Ulfelder KW. Exoglycosidase matrix-mediated sequencing of a complex glycan pool by capillary electrophoresis. J Chromatogr A, 1997,781: 547～554

41 Hu GF. Fluorophore-assisted carbohydrate electrophoresis technology and applications. J Chromatogr A, 1995,705: 89～103

42 Jadach J, Turner GA. An ultrasensitive technique for the analysis of glycoproteins using lecting blotting with enhanced chemiluminescence. Anal Biochem, 1993,212: 293～295

43 Bayer EA, Ben-Hur H, Wilchek M. Enzyme-based detection of glycoproteins on blot transfers using avidin-biotin technology. Anal Biochem, 1987,161: 123～131

44 Che FY, Song JF, Yu LR *et al*. Fluorescent staining of glycoproteins on polyvinylidene difluoride membrane with 8-aminonaphthalene-1, 3, 6-trisulfonate. BioTechniques, 2001,15: 1272～1278

45 Che FY, Song JF, Zeng R *et al*. Analysis of 8-aminonaphthalene-1, 3, 6-trisulfonate-derivatized oligosaccharides by capillary electrophoresis-electrospray ionization quadrupole ion trap mass spectrometry. J Chromatogr A, 1999,858: 229～238

46 James DC, Freedman RB, Hoare M *et al*. High-resolution separation of recombinant human interferon-gamma glycoforms by micellar electrokinetic capillary chromatography. Anal Biochem, 1994,222: 315～322

47 Kilar F, Hjerten S. Separation of the human transferrin isoforms by carrierfree high-performance zone electrophoresis and isoelectric focusing. J Chromatogr, 1989,480: 351～357

48 Tran A, Park S, Lisi PJ *et al*. Separation of carbohydrate-mediated microheterogeneity of recombinant human erythropoietin by free solution capillary electrophoresis. Effects of pH, buffer type and organic additives. J Chromatogr, 1991, 542: 459～471

49 Watson E, Yao F. Capillary electrophoresis separation of human recombinant erythropoietin(rHuEPO)glycoforms. Anal Biochem, 1993, 210: 389～393

50 Cifuentes A, Moreno-Arribas MV, de Frutos M *et al*. Capillary isoelectric focusing of erythropoietin glycoforms and its comparison with flat-bed isoelectric focusing and capillary zone electrophoresis. J Chromatogr A, 1999,830: 453～463

51 Landers JP, Oda RP, Madden BJ *et al*. High-performance capillary electrophoresis of glycoproteins: the use of modifiers of electroosmotic flow for analysis of microheterogeneity. Anal Biochem, 1992, 205: 115～124

52 陈义.糖蛋白微观不均一组分的毛细管电泳分离. 中国科学(B辑), 1998,28: 78～84

53 Righetti PG, Gelfi C, Verzola B *et al*. The state of the art of dynamic coatings. Electrophoresis, 2001,22: 603～611

54 Horvath J, Dolnik V. Polymer wall coatings for capillary electrophoresis. Electrophoresis, 2001,22: 644～655

55 Dolnik V, Hutterer KM. Capillary electrophoresis of proteins 1999～2001.Electrophoresis, 2001,22: 4163～4178

56 Giordano BC, Muza M, Trout A *et al*. Dynamically-coated capillaries allow for capillary electrophoretic resolution of transferrin sialoforms via direct analysis of human serum. J Chromatogr Biomed Sci Appl, 2000,742: 79～89

57 Trout AL, Prasad R, Coffin D *et al*. Direct capillary electrophoresis detection of carbohydrate-deficient transferring in neat serum. Electrophoresis, 2000,21: 2376～2383

58 Taverna M, Baillet A, Biou D *et al*. Analysis of carbohydrate-mediated heterogeneity and characterization of N-linked oligosaccharides of glycoproteins by high performance capillary electrophoresis. Electrophoresis, 1992,13: 359～366

59 Moorhouse KG, Eusebio CA, Hunt G *et al*. Rapid one-step capillary isoelectric focusing method to monitor charged glycoforms of recombinant human tissue-type plasminogen activator. J Chromatogr A, 1995,717: 61～69

60 Yim K, Abrams J, Hsu A. Capillary zone electrophoretic resolution of recombinant human bone morphogenetic protein 2 glycoforms. An investigation into the separation mechanisms for an exquisite separation. J Chromatogr A, 1995,716: 401～412

61 Che FY, Song JF, Wang KY *et al*. Comparative study on the distribution of ovalbumin glycoforms by capillary electrophoresis. J Chromatogr A, 1999,849: 599～608
62 车发云，邵晓霞，夏其昌. 高效液相色谱-电喷雾四极杆离子阱质谱鉴定蛋白质磷酸化位点. 中国科学(C辑)，2000,30: 421～427
63 Annan RS, Huddleston MJ, Verma R, *et al*. A multidimensional electrospray MS-based approach to phosphopeptide mapping. Anal Chem, 2001,73: 394～404
64 Ma Y, Lu Y, Zeng H *et al*. Characterization of phosphopeptides from protein digests using matrix-assisted laser desorption/ionization time-of-flight mass spectrometry and nanoelectrospry quadrupole time-of-flight mass spectrometry. Rapid Commun Mass Spectrum, 2001,15: 1693～1700
65 Liu BF, Zhang L, Lu YT. Characterization of phosphorylation of a novel protein kinase in rice cell by capillary electrophoresis. J Chromatorgr A, 2001,918: 401～409
66 Honda S, Okeda J, Iwanaga H *et al*. Ultramicroanalysis of reducing carbohydrates by capillary electrophoresis with laser-induced fluorescence detection as 7-nitro-2, 1, 3-benzoxadiazole-tagged *N*-methylglycamine derivatives. Anal Biochem, 2000,286: 99～111
67 Lacher NA, Garrison KE, Martin RS *et al*. Microchip capillary electrophoresis/electrochemistry. Electrophoresis, 2001,22: 2526～2536
68 Kricka LJ. Microchips, microarrays, biochips and nanochips: personal laboratories for the 21st century. Clin Chim Acta, 2001,307: 219～223
69 Regnier FE, He B, Lin S *et al*. Chromatography and electrophoresis on chips: critical elements of future integrated, microfluidic analytical systems for life science. Trends Biotechnol,1999,17: 101～106
70 Suzuki S, Shimotsu N, Honda S *et al*. Rapid analysis of amino sugars by microchip electrophoresis with laser-induced fluorescence detection. Electrophoresis, 2001,22: 4023～4031
71 Wen J, Lin Y, Xiang F *et al*. Microfabricated isoelectric focusing device for direct electrospray ionization-mass spectrometry. Electrophoresis, 2000,21: 191～197
72 Li JJ, Wang C, Kelly JF *et al*. Rapid and sensitive separation of trace level protein digests using microfabricated devices coupled to a quadrupole-time-of flight mass spectrometer. Electrophoresis, 2000,21: 198～210

下　　篇

蛋　白　质　组　学

第十二章　导　论

2001年人类基因组框架图的发表使大众对科学家将揭示人类的遗传奥秘充满了幻想和信心。但是,真实的图景远不像普通人想像的那样简单。正如国际人类基因组计划(human genome project,HGP)主要负责人 Francis Collins 在2001年6月一次会上所强调的:真正的竞赛才刚刚开始。竞赛的一个核心内容,就是要研究基因的产物——蛋白质。在 Nature 杂志2001年2月发布人类基因组框架图的消息的同一版面上,编辑也同时登载了一条关于人类蛋白质组组织(human proteome organization,HUPO)成立的消息,标题就叫"现在是蛋白质组了"[1]。2001年的 Science 杂志把蛋白质组学列为2002年六大研究热点之一,其"热度"仅次于干细胞研究[2]。2003年3月 Nature 杂志上的一系列蛋白质组学的综述文章更是把蛋白质组学研究推到了令人瞩目的地位[3~9]。蛋白质组学的受关注程度如今已令人刮目相看。

一、蛋白质组学研究的历史和背景

随着人类基因组计划的实施和推进,生命科学研究已进入了后基因组时代。在这个时代,生命科学的主要研究对象是功能基因组学,包括结构基因组研究和蛋白质组研究等。尽管现在已有多个物种的基因组被测序,但在这些基因组中通常有一半以上基因的功能是未知的。目前功能基因组中所采用的策略,如基因芯片、基因表达序列分析(serial analysis of gene expression, SAGE)等,都是从细胞中 mRNA 的角度来考虑的,其前提是细胞中 mRNA 的水平反映了蛋白质表达的水平。但事实并不完全如此,从 mRNA 角度考虑,实际上仅包括了转录水平调控,并不能全面代表蛋白质表达水平。蛋白质复杂的翻译后修饰、蛋白质的亚细胞定位或迁移、蛋白质和蛋白质相互作用等则几乎无法从 mRNA 水平来判断。毋庸置疑,蛋白质是生理功能的执行者,是生命现象的直接体现者,对蛋白质结构和功能的研究将直接阐明生命在生理或病理条件下的变化机制。蛋白质本身的存在形式和活动规律,如翻译后修饰、蛋白质间相互作用以及蛋白质构象等问题,仍依赖于直接对蛋白质的研究来解决。虽然蛋白质的可变性和多样性等特殊性质导致了蛋白质研究技术远比核酸技术要复杂和困难得多,但正是这些特性参与和影响着整个生命过程。

传统的对单个蛋白质进行研究的方式已无法满足后基因组时代的要求。这是因为:①生命现象的发生往往是多因素影响的,必然涉及到多个蛋白质;②多个蛋白质的参与是交织成网络的,或平行发生,或呈级联因果;③在执行生理功能时蛋白质的表现是多样的、动态的,并不像基因组那样基本固定不变。因此,要对生命的复杂活动有全面和深入的认识,必然要在整体、动态、网络的水平上对蛋白质进行研究。因此,在20世纪90年代中期,国际上产生了一门新兴学科——蛋白质组学(proteomics),它是以细胞内全部蛋白质的存在及其活动方式为研究对象。可以说蛋白质组研究的开展不仅是生命科学研究进入后基因组时代的里程碑,也是后基因组时代生命科学研究的核心内容之一。

虽然第一次提出蛋白质组概念是在 1994 年,但相关研究可以追溯到 20 世纪 90 年代中期甚至更早,尤其是 80 年代初,在基因组计划提出之前,就有人提出过类似的蛋白质组计划,当时称为 Human Protein Index 计划,旨在分析细胞内的所有蛋白质。但由于种种原因,这一计划被搁浅。20 世纪 90 年代初期,各种技术已比较成熟,在这样的背景下,经过各国科学家的讨论,才提出蛋白质组这一概念。

国际上蛋白质组研究的进展十分迅速,不论基础理论还是技术方法,都在不断进步和完善。多种细胞的蛋白质组数据库已经建立,相应的国际互联网站也层出不穷。1996 年,澳大利亚建立了世界上第一个蛋白质组研究中心,即 Australia Proteome Analysis Facility (APAF)。丹麦、加拿大、日本也先后成立了蛋白质组研究中心。在美国,各大药厂和公司在巨大财力的支持下,也纷纷加入蛋白质组的研究阵容。在瑞士成立的 GeneProt 公司,是由以蛋白质组数据库“SWISSPROT” 著称的蛋白质组研究人员成立的,以应用蛋白质组技术开发新药物靶标为目的,建立了配备有上百台质谱仪的高通量技术平台。而当年提出 Human Protein Index 的美国科学家 Norman G. Anderson 也成立了类似的蛋白质组学公司,继续其多年未实现的梦想。2001 年 4 月,在美国成立了国际人类蛋白质组组织(Human Proteome Organization, HUPO),随后欧洲、亚太地区都成立了区域性蛋白质组研究组织,试图通过合作的方式,融合各方面的力量,完成人类蛋白质组计划(Human Proteome Project),并在 2002 年和 2003 年陆续启动了人类血浆、肝和脑的蛋白质组研究国际合作计划(参见 www. hupo. org)。

二、蛋白质组学的特点与难点

根据对国际生命科学文献数据库“Medline”的检索,基因组的研究论文从 1990～2001 年的时间里,头 6 年基本保持不变,只是在 1996 年以后才以大约每年翻一番的速度开始增长。而冠有“蛋白质组”的研究文献,1997 年是 1 篇,到 2003 年已迅速发展至 2 700 多篇。这一结果表明,蛋白质组学迅速为科学界所接受并受到重视。下面笔者将从基因组与蛋白质组比较的角度来谈谈为什么要研究蛋白质组以及蛋白质组学的特点和难点。

(一) 同一性与多样性

基因组作为遗传信息的载体,其最主要的特征就是同一性。对于单细胞生物而言,不论在什么样的生长条件下,其基因组始终是不变的。对于多细胞生物来说,同一个个体的基因组不论是在不同的发育阶段或不同种类的细胞里都是同样的。知道了个体内某一细胞内的基因组就等于知道了该个体所有细胞的基因组。然而,对于蛋白质组而言,由于蛋白质是生命活动的主要执行者,因此,对于不同类型的细胞或同一个细胞在不同的活动状态下,蛋白质组的构成是不一样的。

蛋白质组的这种多样性使得我们对蛋白质组的定义出现了问题。目前蛋白质组比较公认的定义是,一个基因组所有基因所表达的全部蛋白质。这种定义在文字上可以理解,但在实际中却难以运用。任何一种基因组,都是由不编码蛋白质的核苷酸序列和编码蛋白质的核苷酸序列——基因所组成。然而,除了已经知道有蛋白质产物的基因以外,其余的“基因”只不过是根据“起始密码”和“终止密码”序列所确定的“可读框”(open reading

frame, ORF)。一般说来,一个基因组内的基因数目通常是指 ORF 的数目。基因组研究的实践告诉我们,要想准确无误的肯定一个基因组内的 ORF 数目,并不是一件容易的事。虽然生物信息学已经有了长足的进步,但是直接从基因组的序列进行基因注释(annotation)还是很困难的。不久前人们发现,对一个简单的细菌(*Mycoplasma genitalium*)基因组的 340 个基因的注释,其错注的比率至少是 8%[10]。对于那些高等生物复杂的基因组,尤其是存在 RNA 剪切的基因组,要想确定其 ORF 的数目就更成问题了。

显然,最能确定 ORF 的数量应该是通过研究蛋白质组来进行。据最新统计,人类基因组拥有的基因数目为 3 万～4 万个。如果我们能够把人的 252 种细胞内的全部蛋白质都鉴定出来,那么就有可能真正知道人类基因组的所有基因。但是这样一来,基因组和蛋白质组进入了"循环定义"之中:蛋白质组是以基因组拥有的所有基因的表达产物来构成,而所有基因的确定又必须通过蛋白质组来给予肯定。当然,蛋白质是最终的尺度。确定了一个蛋白质就肯定了一个基因,但确定了一个 ORF 却并不能肯定它会有对应的蛋白质产物。

即使基因注释或 ORF 数目的问题我们能够解决,这些 ORF 是否与蛋白质存在一一对应关系也仍是一个问题。一方面,人们已经发现有许多"拟基因"(pseudogene)的存在,这些拟基因有和真基因相同的 ORF,但却从不表达。另一方面,由于存在 RNA 水平上遗传信息的加工——mRNA 编辑(RNA editing)和蛋白质水平上遗传信息的加工——蛋白质剪切(protein splicing),许多蛋白质很难找到直接对应的 ORF。所以,要想找出一个生物体基因组的全部基因和相应的全部蛋白质是一个非常困难的任务。

(二) 有限与无限

对于基因组研究而言,核心任务是测定所有的 DNA 碱基序列。例如,芽殖酵母基因组有 1 200 万碱基对,人类基因组的最新统计为 32 亿个碱基对。要测定的基因组不论大小,其核苷酸的数量是明确的。然而,对蛋白质组来说,蛋白质的种类究竟有多少就很难说了。上面说过,蛋白质组可以被定义为基因组基因表达的所有蛋白质,但这一定义没有考虑蛋白质的修饰。细胞内的大部分蛋白质通常都被进行过翻译后修饰,主要有磷酸化、糖基化、酰基化等。修饰过的蛋白质具有其特定的物理化学性质和生物学功能。对蛋白质修饰的研究已构成了蛋白质组研究的一个重要部分,由此还形成了蛋白质组学的一个分支:修饰蛋白质组学[11]。如果把一个修饰蛋白视为一种新的蛋白质,那么蛋白质组的蛋白质数量将远远大于相应的基因组的基因数量。在这个意义上,人们估计人类蛋白质组的蛋白质的种类为 20 万～200 万[12]。显而易见,对蛋白质组的蛋白质数量的估计是非常模糊的。

从蛋白质修饰的角度来看,不仅仅是蛋白质组的蛋白质种类大大增加,更重要的是,由于不存在度量蛋白质修饰的尺度,人们也许永远不能像确定基因组核苷酸序列那样肯定地把握生物体内蛋白质组的蛋白质数量。如果说表达的蛋白质的种类可以根据基因的数目来确定,那么修饰形成的蛋白质种类只有依靠对蛋白质的直接研究来决定,而这种研究是没有止境的,因为蛋白质的修饰和生命的活动密切相关。从这种意义上来说,对基因组核苷酸序列的测定是一种"有限"的工作,而对蛋白质组的蛋白质种类的确定则是一种"无限"的工作。

(三) 静态与动态

一个个体的基因组自个体诞生到死亡,始终保持不变。而作为新陈代谢的主要执行者的蛋白质组,在个体的生命活动中却总是变动不停。人们可以通过确定变化的蛋白质来理解其功能。因此,蛋白质组研究的一个重要任务是,找出蛋白质组里发生了变化的蛋白质。生命活动中的蛋白质可以大致分为两类:一类是比较稳定的,如负责细胞各种结构组成的蛋白质。另一类则是随细胞的活动和状态发生量或质变化的蛋白质,如负责信号转导或细胞活动调控的蛋白质。在这种意义上,人们把以研究蛋白质变化为主的蛋白质组学称为功能蛋白质组学(functional proteomics)。它主要是比较在不同生长状态下或病理情况下的同一种细胞或组织的蛋白质组内的差异蛋白质。与此相对应,测定一个细胞或组织含有的全部蛋白质种类的研究就是通常所说的蛋白质组学。例如“Large Scale Proteomics”公司正在从事的“人类蛋白质索引”(Human Protein Index, HPI)项目,就是要建立一个人类所有蛋白质的数据库;据说目前已从157种不同的组织中得到11.5万种蛋白质[13]。

这种对蛋白质动态的研究又引出了蛋白质组学的另一个问题,即怎样区分差异蛋白质是属于生物学意义上的还是实验意义上的。因为蛋白质的稳定性要比核酸差得多,在制备的过程中可能会发生降解或丢失。此外,差异蛋白质的区分通常是通过蛋白质双向凝胶电泳来检测,而双向电泳由于自身分辨率的限制,难以测定微量的蛋白质变化。有实验表明,目前的双向电泳技术对酵母细胞中低于1 000个拷贝量的蛋白质是检测不到的[14]。还有一个问题是,对于大多数差异蛋白质而言,有或无这类质的变化的差异蛋白质较少,而在量的增减方面的差异蛋白质则较多。由于双向电泳技术的不稳定性,所以要准确测定蛋白质量的差异是一个困难的任务。因此,从动态的角度来研究蛋白质组是当前蛋白质组学的一个挑战。

尽管测定差异蛋白质的技术难度较大,但功能蛋白质组学在今后一段时间将仍然是蛋白质组研究领域的主要方向。首先,蛋白质的差异和变化是生命活动的重要方式;发现在正常和病理状态下出现差异的蛋白质常常是认识和防治疾病的关键。因此大多数研究者仅仅关心寻找蛋白质的差异表达谱。其次,鉴定出一个细胞或组织中的所有蛋白质是一项费时费钱的工作,需要大量的人力和物力。相比之下,测定差异蛋白质的任务需要的投入就小得多。

(四) 时间与空间

DNA通常位于细胞核内,且保持稳定,因此测定基因组的DNA序列是不受时空的影响。对于转录的mRNA来说,时间是主要的参考因素,在发育的不同阶段或细胞的不同活动时期,mRNA的表达是不一样的。因此,在研究转录组或基因芯片时必须考虑到时间,但通常不需要考虑空间的影响。而在蛋白质组的研究中,不仅要考虑时间的因素,更要考虑空间的影响。首先,不同的的蛋白质分布在细胞的不同部位。它们的功能与其空间定位密切相关。要想真正了解蛋白质的功能,必须要知道蛋白质所处的空间位置。其次,许多蛋白质在细胞里不是静止不动的,它们在细胞里常常通过在不同的亚细胞环境里的运动发挥作用。例如,在细胞周期的调控过程中,磷酸酯酶Cdc25在细胞核内对蛋白激

酶 Cdc2 进行去磷酸化，从而激活 Cdc2，使细胞进入分裂期。如果 DNA 损伤发生，Cdc25 会被一种 14-3-3 蛋白从细胞核内转运到胞质内，导致核内的 Cdc2 保持磷酸化和失活状态，使细胞停滞在 G_2 期[15]。细胞的信号传导和转录调控也常依赖于蛋白质在空间位置的变化和运动。因此，蛋白质组学中又派生了一个与空间紧密相关的新研究领域——亚细胞蛋白质组学[16]，这种亚细胞蛋白质组可能是细胞器蛋白质组，如高尔基体蛋白质组[17]；也可能是比细胞器还要小的组分——如核膜[18]。对亚细胞蛋白质组的研究将有助于对亚细胞结构和功能的理解。此外，值得一提的是，分级分离技术对于研究细胞的整体蛋白质组也是很有用的。目前双向电泳技术能分离的蛋白质数量最多是几千种，而一个细胞内的蛋白质数量通常会超过万种，尤其是高等真核生物的细胞，可能有数万种蛋白质。因此，用双向电泳技术分离全细胞的蛋白质组分的效果很不理想。如果利用细胞的分级分离技术，人们可以把细胞不同组分的蛋白质谱分别进行鉴定，然后再进行“组装”，从而得到细胞完整的蛋白质谱。

对于亚细胞蛋白质组学来说，关键技术是对细胞进行恰当的分级分离。由此，蛋白质组研究又面临着一个新的挑战：如何获得完整的亚细胞成分。在这种分级分离过程中，既要避免待研究的亚细胞组成蛋白质的丢失，又要避免非组成蛋白质的“污染”。另外，如何区别亚细胞组成蛋白质和由于生理作用而在细胞不同区域进行运动的蛋白质也是亚细胞蛋白质组的一个技术难点。例如，负责转运蛋白质的高尔基体，既有组成蛋白质，也有待输送的蛋白质，在研究高尔基体蛋白质组时应该加以区分。

（五）孤立行为与相互作用

基因组的基因表达的各种 mRNA 是彼此孤立的，互不干扰。但是 mRNA 的产物——蛋白质正好相反，它们彼此间有着广泛的相互作用。也许可以说，不存在不与其他蛋白质发生作用的“孤立蛋白质”。蛋白质的功能的实现离不开蛋白质与蛋白质或蛋白质与其他生物大分子之间的相互作用。所以，蛋白质组研究的一个主要内容是研究大规模的蛋白质的相互作用。科学家为此专门发明了一个新的名词——“相互作用组”（interactome）[19]。

相互作用组研究可以分为两类。第一类是研究蛋白质相互作用的网络。细胞内的许多活动如信号转导等都是通过一个复杂而广泛的蛋白质相互作用网络来实现的。不久前，科学家通过大规模双杂交的手段，利用 27 种蛋白质对线虫的基因组进行了筛选，得到了参与线虫生殖器发育的 124 种发生相互作用的蛋白质，其中大约有 100 种是未知功能的蛋白质[20]。对这一类型的研究来说，除了大规模双杂交技术外，人们还使用了噬菌体显示技术（phage-display）和蛋白质芯片技术[22]。对于大规模蛋白质相互作用分析而言，最突出的技术问题是虚假信号比较多，如双杂交技术容易出现假阳性，而噬菌体显示技术则容易产生假阴性。显然，如果把不同技术结合起来，可以把结果进行互补，从而更准确地描绘出蛋白质的相互作用网络。

相互作用组的另一类研究是蛋白质复合体组成的分析。蛋白质复合体通常可以分为两种。一种是结构型的蛋白质复合体，如核孔复合体；这一类通常比较稳定。另一种则是功能型蛋白质复合体，例如，负责转录的转录蛋白复合体或者负责 DNA 复制的复制蛋白复合体；这类复合体只有在执行功能时才聚合在一起，因此很不稳定。对于结构型蛋白质

复合体，通常采用分级分离的方法进行富集，然后再用经典的蛋白质组研究方法如电泳和质谱进行研究。但对于功能型蛋白质复合体，现有的方法就很难胜任了。最近，研究者发明了一种新的分离方法，可能有助于对这类蛋白质复合体的研究。首先是构造能够表达含有特定标签肽段(tag)的复合蛋白质的质粒，这些质粒表达的蛋白质在正常生理状态下与细胞内能够与其发生相互作用的其他种类蛋白质形成复合物；然后通过对标签肽段有专一亲和性的亲和分离柱进行提取，最后再通过电泳和质谱对这些蛋白质复合物的组分进行分析[23,24]。发展对蛋白质复合体的研究技术将是今后蛋白质组学的一项重要任务。

（六）单一手段与多种技术

对于研究核酸和蛋白质这类生物大分子来说，测定其序列是首要任务。在 20 世纪 50 年代分子生物学刚刚诞生之际，科学家就发明了测定蛋白质氨基酸序列的方法[25]。在 70 年代研究者又发明了测定 DNA 核苷酸序列的化学法和酶法[26,27]。随后人们发现，测定 DNA 的核苷酸序列要远比测定蛋白质的氨基酸序列容易得多。因此，在 80 年代和 90 年代，测定 DNA 序列成为分子生物学研究的一个主要手段，而蛋白质序列的测定则很少人问津。即使需要知道某种蛋白质的序列，研究者也尽可能试图先得到其编码的基因或 cDNA，测定其核苷酸序列，最后再推出对应的蛋白质序列。但是，对于蛋白质组研究而言，测序工作不再是一个可以绕开的问题。生物质谱技术的出现，为蛋白质序列测定提供了一个最主要的手段。这也就是为什么质谱技术现已成为蛋白质组研究的核心技术的原因。

前面已讨论过，基因组既无量的变化，也没有质的变化。在基因组研究中，DNA 测序技术是一个最基本和最主要的工具。这是因为基因组的均一性和简单性使得一种单一的技术就能胜任基因组研究任务。但是，在蛋白质组研究中，需要的研究技术就远远不止一种，并且技术的难度也远远大于基因组研究技术。

蛋白质组研究技术可以简单地分为两大类。第一类是蛋白质组的分离技术。前面所说过的双向电泳技术、细胞分级分离技术、亲和层析技术都是蛋白质组常用的分离技术。蛋白质组分离有这样几个难点。首先是量的问题，不同的蛋白质数量在细胞内可以有很大的差别。例如，在酵母细胞里，有些细胞周期调控蛋白不到 100 个分子，而糖基酶可能有 200 万个分子[28]。估计蛋白质量的动态差别可达到 10^6 数量级。其次是质的问题，蛋白质是由 20 种化学性质各异的氨基酸所组成，因此不同的蛋白质的物理化学性质差别很大。有些蛋白质易溶于极性溶剂，有些蛋白质则难溶于极性溶剂；有些蛋白质较稳定，有些蛋白质则易降解。目前膜蛋白的提取仍很不理想。

第二类是蛋白质组的鉴定技术，其核心是质谱技术。蛋白质芯片技术、噬菌体显示技术和大规模双杂交方法等也属于鉴定技术。现有鉴定技术面临的一个最大问题是，它们都依赖于已知的蛋白质或基因的序列作为检测的基础，通过比较来确定待测定的蛋白质。对于在已有数据库中没有记载的全新蛋白质进行“从头测定”(*de novo* identification)是一个很困难的任务，一般是利用经典的 Edman 降解法进行 N 端序列测定或 C 端化学降解法测定 C 端序列。

简而言之，蛋白质组研究对技术的依赖性和要求远远超过基因组学。目前，蛋白质组学的研究技术还有很多不完善之处，许多新技术正在研发之中。因此，蛋白质组学的发展

是受技术限制的，也是受技术推动的。

（七）互补与互助

自从20世纪中叶分子生物学因DNA结构和蛋白质结构的解析而诞生以来，现代生命科学的主旋律一直是围绕着核酸与蛋白质展开。在后基因组时代，蛋白质组研究和基因组研究依然是形影相随的两个重要领域。它们之间既为互相补充又能互相帮助。

当基因组测序工作结束之后，下一个主要任务是要进行基因注释(annotation)，找出所有可能的基因或ORF。前面曾经说过，一般情况下注释的出错率是比较高的[10]。但是，如果有蛋白质组的数据作参考，可以大大提高注释的正确率。此外，由于高等真核生物个体的基因大多是非连续性的，即由编码氨基酸的外显子和非编码氨基酸的内含子所组成，因此确定完整的ORF也非易事。显然，利用蛋白质的序列来推测和确定ORF的全长也是很有帮助的。

mRNA是介于基因和蛋白质之间的中间产物。因为它既是基因的直接产物，又比蛋白质要容易分析，所以，研究mRNA的表达模式也是了解基因组和蛋白质组的一个重要途径。为此专门形成了一个新的研究领域——转录组(transcriptome)。研究转录组的主要手段是基因芯片技术，此外如SAGE技术也可用于转录组研究[29]。有报道认为，可以从mRNA的转录水平去测定相应的蛋白质的含量。在酵母细胞里，一个mRNA分子大约能翻译出4 000个蛋白质分子[28]。

蛋白质组的许多工作也离不开对基因组的研究。这一点在蛋白质相互作用的研究中尤为突出。双杂交技术、噬菌体显示技术等就是以基因组数据和技术为基础的。在质谱测定蛋白质的过程中，离不开对已知基因组的基因数据的比较。有人把蛋白质组研究比做一种双向研究：经典的蛋白质组学是从蛋白质组到转录组再到基因组，而“反向蛋白质组学”(reverse proteomics)则是从基因组到转录组再到蛋白质组[30]。

三、蛋白质组学发展趋势

蛋白质组学一经出现，就有两种研究策略。一种可称为“竭泽法”，即采用高通量的蛋白质组研究技术分析生物体内尽可能多乃至接近所有的蛋白质，这种观点从大规模、系统性的角度来看待蛋白质组学，也更符合蛋白质组学的本质。但是，由于蛋白质表达随空间和时间不断变化，要分析生物体内所有的蛋白质是一个难以实现的目标。另一种策略可称为“功能法”，即研究不同时期细胞蛋白质组成的变化，如蛋白质在不同环境下的差异表达，以发现有差异的蛋白质种类为主要目标。这种观点更倾向于把蛋白质组学作为研究生命现象的手段和方法。

早期蛋白质组学的研究范围主要是指蛋白质的表达模式(expression profile)，随着学科的发展，蛋白质组学的研究范围也在不断完善和扩充。蛋白质翻译后修饰研究已成为蛋白质组研究中的重要部分。蛋白质-蛋白质相互作用的研究也已被纳入蛋白质组学的研究范畴。而蛋白质高级结构的解析即传统的结构生物学，虽也有人试图将其纳入蛋白质组学研究范围，但目前仍独树一帜。

在基础研究方面，近两年来蛋白质组研究技术已被应用到各种生命科学领域，如细胞

生物学、神经生物学等。在研究对象上，覆盖了原核微生物、真核微生物、植物和动物等范围，涉及到各种重要的生物学现象，如信号转导、细胞分化、蛋白质折叠等。在未来的发展中，蛋白质组学的研究领域将更加广泛。

在应用研究方面，蛋白质组学将成为寻找疾病分子标记和药物靶标最有效的方法之一。在对癌症、早老性痴呆等人类重大疾病的临床诊断和治疗方面蛋白质组技术也有十分诱人的前景，目前国际上许多大型药物公司正投入大量的人力和物力进行蛋白质组学方面的应用性研究。

在技术发展方面，蛋白质组学的研究方法将出现多种技术并存、各有优势和局限的局面，而难以像基因组研究一样形成比较一致的方法。除了发展新方法外，更强调各种方法间的整合和互补，以适应不同蛋白质的不同特征。另外，蛋白质组学与其他学科的交叉也将日益显著和重要，这种交叉是新技术、新方法的活水之源，更重要的是，蛋白质组学与其他大规模科学如基因组学、生物信息学等领域的交叉，所呈现出的系统生物学(systems biology)研究模式，将成为未来生命科学最令人激动的新前沿。

(吴家睿　曾　嵘)

参考文献

1 Abbott A. And now for the proteome. Nature, 2001,409: 747

2 Service RF. High-speed biologists search for gold in proteins. Science,2001,294:2074～2077

3 Marte B. Proteomics. Nature,2003,422: 191

4 Tyers M, Mann M. From genomics to proteomics. Nature, 2003,422: 193 ～197

5 Aebersold R, Mann M. Mass spectrometry-based proteomics. Nature, 2003,422: 198～207

6 Phizicky E, Bastiaens PI, Zhu H *et al*. Protein analysis on a proteomic scale. Nature, 2003,422: 208～215

7 Sali A, Glaeser R, Earnest T *et al*. From words to literature in structural proteomics. Nature, 2003,422: 216～225

8 Hanash S. Disease proteomics. Nature, 2003, 422: 226～232

9 Boguski MS, McIntosh MW. Biomedical informatics for proteomics. Nature, 2003,422: 233 ～237

10 Brenner, S. Errors in genome annotation. Trends Genet., 1999,15:132～133

11 Oda Y, Nagasu T, Chait BT. Enrichment analysis of phosphorylated proteins as a tool for probing the phosphoproteome. Nat Biotechnol, 2001,19: 379～382

12 Service RF. High-speed biologists search for gold in proteins. Science, 2001,294: 2074～2077

13 Abbott A. Human proteome index launched. Nature, 2001,409: 3

14 Gygi SP, Rochon Y, Franza BR *et al*. Correlation between protein and mRNA abundance in yeast. Mol Cell Biol, 1999,19: 1720～1730

15 Lopez-Girona A, Furnari B, Mondesert O *et al*. Nuclear localization of Cdc25 regulated by DNA damage and 14-3-3 protein. Nature, 1999,397: 172～175

16 Jung E, Heller M, Sanchez JC *et al*. Proteomics meets cell biology: the establishment of subcellular proteomes. Electrophoresis, 2000,21: 3369～3377

17 Taylor RS, Wu CC, Hays LG *et al*. Proteomics of rat liver Golgi complex: minor proteins are identified through sequential fractionation. Electrophoresis, 2000,21: 3441～3459

18 Dreger M, Bengtsson L, Schoneberg T *et al*. Nuclear envelope proteomics: novel integral membrane proteins of the inner nuclear membrane. Proc Natl Acad Sci USA., 2001,98: 11943～11948

19 Gerstein M, Lan N, Jansen R. Integrating interactomes. Science, 2002,295: 284～287

20 Walhout AJ, Walhout AJ, Sordella R, *et al*. Protein interaction mapping in *C. elegans* using proteins involved in vul-

val development. Science, 2000,287: 116～122
21 Tong AH, Evangelista M, Parsons AB *et al*. Systematic genetic analysis with ordered arrays of yeast deletion mutants. Science, 2001,294: 2364～2368
22 Zhu H, Bilgin M, Bangham R *et al*. Global analysis of protein activities using proteome chips. Science, 2001, 293: 2101～2105
23 Gavin AC, Bosche M, Krause R *et al*. Functional organization of the yeast proteome by systematic analysis of protein complexes. Nature, 2002,415: 141～147
24 Ho Y, Gruhler A, Heilbut A *et al*. Systematic identification of protein complexes in *Saccharomyces cerevisiae* by mass spectrometry. Nature, 2002,415: 180～183
25 Edman P. Method for determination of the amino acid sequence in peptides. Acta Chem Scurd., 1950.4: 283
26 Maxam AM, Gilbert W. A new method of sequencing DNA. Proc Natl. Acad Sci USA, 1977,74: 560～564
27 Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain terminating inhibitors. Proc Natl Acad Sci USA, 1977,74: 5463～5467
28 Futcher B, Latter GI, Monardo P *et al*. A sampling of the yeast proteome. Mol Cell Biol, 1999,19: 7357～7368
29 Velculescu VE, Velculescu VE, Zhang L *et al*. Characterization of the yeast transcriptome. Cell, 1997,88: 243～251
30 Walhout AJ, Vidal M. Protein interaction maps for model organisms. Nat Rev Mol Cell Biol, 2001,2:55～62

第十三章　样品的全息制备

蛋白质样品制备是蛋白质组研究的第一步,无论后续采用怎样的分离或鉴定手段,样品制备都是关键步骤。从蛋白质组大规模研究的角度而言,要求样品制备尽可能获得所有的蛋白质,但由于蛋白质种类多、丰度不一和物化特性多样等,要达到真正的全息制备是有难度的。当然,样品制备的方法还必须与后续的分离或鉴定方法相匹配,如采用双向凝胶电泳,需谨慎使用SDS抽提蛋白质,如果以液相色谱作为分离手段,太多的去污剂和两性电解质不是一个好的选择。更重要的是,样品的来源千差万别,针对不同来源的样品(如细胞样品、动物组织样品、植物样品、体液等),需要采取不同的蛋白质抽提方法。本章就常规样品的制备和不同来源蛋白质的样品制备进行详细阐述。

第一节　双向凝胶电泳常规样品制备及其改进

一、概　　述

目前蛋白质组研究技术发展很快,新的研究方法和研究手段不断涌现,如毛细管等电聚焦[1]、多维色谱[2]、分子扫描仪和微流芯片[3]等。而双向凝胶电泳(two-dimensional gel electrophoresis)技术是蛋白质组研究中最早的支撑技术之一,自O′Farell[4]等提出以来,已有近30年的历史。由于双向凝胶电泳技术具有高分辨率、高重复性、微量分析和制备等性能,仍然是当今蛋白质组分析研究中的一个强有力的工具。因而针对双向凝胶电泳样品制备方法的探索也是蛋白质组研究的一个重要方面。

目前双向凝胶电泳的双向一般是指第一向为等电聚焦(IEF),根据蛋白质等电点进行分离;第二向为SDS凝胶电泳(SDS-PAGE),根据蛋白质的相对分子质量进行分离。所以,在样品制备过程中,一切影响等电聚焦和凝胶电泳的因素都应该考虑在内。样品制备绝对是一门艺术,方法可以从简单的缓冲液溶解,到应用离液剂、还原剂及去污剂来进行复合物提取,以及到更复杂的顺序提取(sequence extracting)[5,6]和亚细胞分级等。由于实验需求的不同以及样品本身的特殊性,没有一种样品制备方法能够广泛地适用于各种各样的样品。所以,对于每一个有着不同要求的样品,都需要通过大量的实验来摸索最合适的实验条件。有效的可重复的样品制备方法是双向凝胶电泳成功的关键[7~10]。一般来说,一种理想的有效的样品制备方法应该满足以下几点要求。

①在合适盐浓度下,可重复地溶解所有的蛋白质,包括疏水性蛋白质,并使样品中的蛋白质以分离的多肽链形式存在。

②避免溶解性低的蛋白质(如膜蛋白)在等电聚焦时由于溶解度降低而沉淀析出。

③防止蛋白质的化学修饰,包括蛋白质降解、蛋白酶或尿素热分解后所引起的修饰。

④排除核酸、多糖、脂类和其他干扰分子。

⑤获得的目标蛋白质应在可探测水平,有时需要去除高丰度蛋白质和不相关蛋白质。

二、样品裂解液中各种成分分析

为获得最佳的双向电泳分离结果，样品一般使用促溶剂、去污剂、还原剂、IPG 缓冲液和两性电解质等将其变性并充分溶解。然而这些化学物质应满足一些化学和电学要求，如适合 IPG 胶条的 IEF 技术、保持蛋白质最大的溶解性和尽可能低的离子强度(等电聚焦加以高电压时，低离子强度使得电流很低，保证了聚焦的快速和高效)。关于溶解方法更多具体内容在许多相关文章中均有讨论[10~12]。

(一) 离液剂

尿素是最常用的离液剂(chaotropic agent)，它的主要作用是改变或破坏氢键等次级键的结构，使得蛋白质变性并使蛋白酶失活。硫脲和尿素联合使用，可以大大增加蛋白质的溶解性，特别是膜蛋白的溶解性[13,14]。尿素和硫脲有效地破坏了疏水键，防止了聚集作用和二级结构的形成，聚集作用和二级结构的形成会改变蛋白质的迁移性。然而硫脲在水中的溶解性很差，需要高浓度的尿素来助溶，因此在实际应用中，需要高溶解强度时，一般用 2mol/L 硫脲和 5～8mol/L 尿素联合使用。

(二) 还原剂

还原剂(reducing agent)主要是用来断裂蛋白质分子中 Cys 残基之间形成的二硫键，增加蛋白质的溶解性。常用的还原剂有 β-巯基乙醇、DTT 或 DTE(Dithioerythreitol)和三丁基膦(TBP)等。其中 DTT 是使用比较广泛的还原剂，它在 50mmol/L 时能有效地还原大部分的二硫键，但有些蛋白质仍不能完全被还原。但是如果过分提高 DTT 的浓度，由于它的 pK_a 在 8 左右，因而会影响到 pH 梯度。另外，DTT 在碱性 pH 值下会发生去质子化，在等电聚焦时，向阳极发生迁移，使得阴极的 DTT 损耗而不足[15]，导致二硫键复原，蛋白质沉淀。而非离子型还原剂 TBP 则比 DTT 更加有效，能大大增加蛋白质的溶解性，在低摩尔浓度下就能使蛋白质得以还原[16]。

(三) 去污剂

去污剂的作用主要是破坏蛋白质分子之间的疏水相互作用，提高蛋白质的溶解性，防止在等电聚焦时析出。去污剂有离子型、非离子型和兼性离子型等几类。经常使用的去污剂有阴离子去污剂 SDS，非离子去污剂 TritonX-100 和 NP-40，两性离子去污剂 CHAPS 等。原则上去污剂应该是非离子型或者是兼性离子型，这样才不会影响蛋白质迁移到它们各自的等电点位置。离子型去污剂如 SDS 不利于等电聚焦，然而由于 SDS 的缓冲液可抑制蛋白质被内源性蛋白水解酶降解，并提高了蛋白质的溶解性[17]，特别是膜蛋白的溶解性，因此一般只用于样品处理的初级阶段。在 SDS 浓度不高于 0.3%时，对等电聚焦不会产生太大的影响。传统的非离子型去污剂，如 Triton X-100 和 NP-40，以前用得较多，现多数以兼性离子去污剂 CHAPS(3-[(3-Cholamidopropyl)-Dimethylammonio]-1-ProPane Sulfonate)代替。CHAPS 比其他去污剂溶解疏水性氨基酸残基的能力更强，其使用浓度一般在 1%～2%。然而它在高浓度脲中的溶解度比较低，因此不断有新的去污剂涌

现出来,如新的兼性离子去污剂 ASB-14[18](amidosulfobetaine 14),Molloy 等人[19]把有机溶剂抽提的大肠杆菌蛋白质溶于含 ASB-14 的增溶溶液中,发现蛋白质的溶解性比用 SB 3-10 和 CHAPS 好;从 13 个被鉴定的蛋白质中发现 8 个是以前在双向胶上没有发现过的,其中 5 个都是疏水性蛋白质。

(四) 两性电解质载体

在去污剂存在的情况下,某些蛋白质在等电聚焦的时候仍旧容易沉淀在它们的等电点上。为了防止这种现象的出现需要向蛋白质溶液中加入某些盐类以增加蛋白质溶解度,但是与此同时这些盐类又会使聚焦效果不理想。为了解决这个问题,通常在缓冲液中加入两性电解质作为载体。两性电解质能够促进蛋白质的溶解,吸附高浓度尿素在溶液中形成的氰酸盐离子(cyanate ions),离心时还有助于核酸的沉淀。此外,它还可以阻止样品蛋白质和 IPG 胶条中固相化的两性电解质相互作用。一定浓度的两性电解质会提高蛋白质溶解性,在第一向等电聚焦时,提高极性端蛋白质的聚焦效果,如图 13.1 所示。

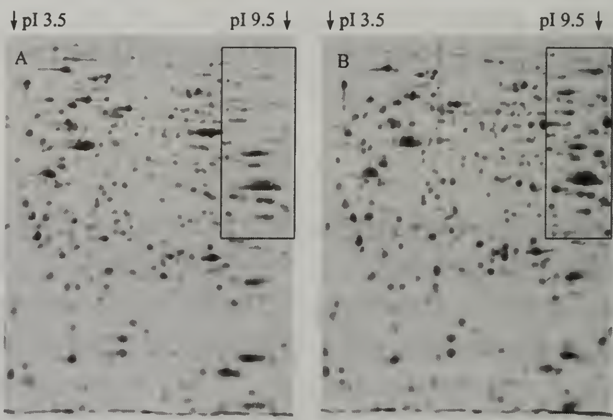


图 13.1 两性电解质对 2-DE 分辨率的影响

样品取自培养的 ECV 304 细胞,A 和 B 上样量均为 10μg,第一向为 7cm pH 3~10 NL IPG 胶条,第二向为 11%的 PAGE。A,两性电解质浓度为 0.5%; B,两性电解质浓度为 2%

以上介绍的是常规使用的各种裂解液成分对样品蛋白质的作用,保证蛋白质达到最大的溶解度并适合双向凝胶电泳。关于裂解液的配方,由于不同实验室有着不同的习惯和需求,配方也千差万别。Görg[15]总结了三种裂解液的配方。

①9mol/L 脲, 1% (*m*/*V*) DTT, 2%~4% (*m*/*V*) CHAPS, 2% (*V*/*V*) 两性电解质, pH3~10 和 10mmol/L Pefabloc 蛋白酶抑制剂。

②7mol/L 脲, 2mol/L 硫脲,1% (*m*/*V*) DTT, 2%~4% (*m*/*V*) CHAPS, 2% (*V*/*V*) 两性电解质, pH3~10 和 10mmol/L Pefabloc 蛋白酶抑制剂。

③热的 SDS 缓冲液:1% (*m*/*V*) SDS,100mmol/L Tris-HCl, pH7.0,该缓冲液裂解样品后至少 3 倍稀释于 1 或 2 的裂解液中。根据目标蛋白质的特性和实验要求,我们同样可以摸索出更合适的裂解液配方,进行有效重复的样品制备。

三、双向凝胶电泳常规样品制备

样品制备时要使用新鲜的样品或者将新鲜样品速冻（液氮或 – 70℃）保存；注意防止污染，佩戴乳胶手套等。整个操作过程在冷库或冰浴中进行，避免蛋白质的降解、修饰。下面以最简单的培养细胞的样品制备为例，介绍操作步骤如下。

①常规细胞培养，待长至 80％左右密度时，在处理前一天更换新鲜培养液。

②吸去培养液，用预冷 PBS 清洗细胞，重复 3 次。

③按 1×10^6 细胞加 50μl 的量往培养皿中加裂解液（8mol/L Urea，4％ CHAPS，40mmol/L Tris，65mmol/L DTT），刮刀收集。

④进行超声波处理，处理时间为 5s，间歇时间为 10s，功率为 100～120W，超声处理至溶液清澈无黏稠物为止。

⑤ 4℃、25 000*g* 离心 1h。

⑥取上清测蛋白质浓度后，分装后 – 70℃ 冰箱中冻存。

四、蛋白质顺序提取法

由于在 2-DE 胶中所能显示出的蛋白质的数量，依赖于细胞中蛋白质的拷贝数、蛋白质的上样量及电泳后染色方法的灵敏度。为了能够富集低拷贝蛋白质，很多实验室采用了样品预分级的方法，如亚细胞分级、亲和分级、蛋白质复合物分级和根据蛋白质溶解性的顺序提取法[5,6]等。另外，窄 pH 范围的 IPG 胶条在一定意义上也可以说是富集了低拷贝蛋白质，提高了低丰度蛋白质在 2-DE 中的分辨率。

这里主要介绍蛋白质顺序提取法，它是根据蛋白质溶解性差异，用具有不同溶解能力的溶解液进行顺序提取，从而提高了样品在 2-DE 中的可检测范围。其具体操作步骤如下所述。

①溶液 A（40mmol/L Tris）提取：将细胞裂解，振荡混匀，12 000r/min 离心 8min，取上清，抽提出细胞中的可溶性蛋白质。

②溶液 B（8mol/L 脲，4％CHAPS，25mmolDTT）提取：在上一步的沉淀物中加入溶液 B，振荡混匀，12 000r/min 离心 8min，取上清。

③溶液 C（5mol/L 脲，2mol/L 硫脲，2％CHAPS，2％SB3-10、2mmol/L TBP）提取：上一步的沉淀物用溶液 A 洗 2 次，加溶液 C，振荡混匀，12 000r/min 离心 8min，取上清。

如有必要也可用更强烈的溶解液进行第四步提取，这样就根据蛋白质溶解性不同，提取出了不同的蛋白质组分，对疏水性很强的蛋白质如膜蛋白和低丰度的蛋白质起了富集作用。顺序提取法是目前获得更多蛋白质点的一种良好策略。应用双向凝胶电泳分离手段，一步提取法一般只能获得 1 000～2 000 个不同的蛋白质点；而使用分级顺序提取方法，对不同组分进行分离，据报道能够获得上万个不同的蛋白质点。在样品量充足的情况下，先进行亚细胞器分级，然后对各个分级组分再进行顺序提取，可以在一定程度上富集溶解度差及低丰度的蛋白质，尽量把细胞中的全部蛋白质提取出来。

五、注意事项

以上介绍的是双向凝胶电泳蛋白质样品制备的一些常规方法。然而这些制备方法并不是十分完善的,相信随着不断的实验探索,会有更理想的制备方法出现。另外,由于每一个样品的特殊性,不同实验中应根据自己的需要对众多的样品制备方法加以选择或者修改,但不管采用什么方法,都应该注意以下问题。

①在合适的盐浓度下,应保持蛋白质的最大溶解性和可重复性。因此,在样品制备过程中可选择合适的增溶试剂(如脲、硫脲等),以减少疏水性蛋白质的损失。

②选择合适的表面活性剂和还原剂,破坏所有非共价结合的蛋白质复合物和共价键二硫键,使其形成一个各自多肽的溶液。

③尽量除去核酸、多糖、脂类等干扰分子。

④防止蛋白质在样品处理过程中的人为修饰,制备过程应在低温下进行,以避免细胞破碎释放出的各种酶类的修饰(建议加入合适的蛋白酶抑制剂);含有脲的试剂不要加热(温度不高于 37℃),避免尿素分解产生的异硫氰酸盐引起氨基甲酰化而导致电荷的不均一性;佩戴乳胶手套,减少来自头发和皮肤的角蛋白污染。

⑤裂解液新鲜配置,分装后置 -80℃ 冰箱中保存,注意不要反复冻融。

⑥样品建议分装成合适的量,然后冷冻干燥或者直接以液体状态置 -80℃ 冰箱中保存,但要注意不要反复冻融。

第二节 培养细胞蛋白质样品的制备

在蛋白质组研究中,我们经常利用细胞培养技术,观察细胞在不同生理条件下的蛋白质动态变化。下面主要介绍关于细胞破碎样品的几种制备方法。

一、细胞破碎的处理方法

细胞破碎主要包括机械和化学方法两种。由于在细胞破碎时会释放出一些蛋白酶及其他引起蛋白质修饰的酶。为了避免这种情况,样品中应加蛋白酶抑制剂,同时尽量将样品直接加入强变性裂解液中,在低温及使用预冷的溶液下进行操作。根据裂解程度的不同,细胞破碎方法又可分为以下两种。

(一)温和的裂解方法

裂解对象主要是一些容易裂解的细胞,如组织培养细胞、血细胞和微生物等。同时温和裂解方法可用于一些亚细胞分级产物,如线粒体、高尔基体和膜等。有时在制备样品的时候需要多种裂解方法综合使用来获得目标蛋白质。如表 13.1 所示为温和裂解方法的应用对象和操作步骤。

表 13.1 温和的裂解方法

细胞破碎方法	应用对象	操作步骤
渗透溶胞法[20]		
非常温和的裂解方法，可用于进一步进行亚细胞分级	血细胞 组织培养细胞	将细胞悬浮于渗透溶胞溶液中，细胞溶胀而释放出细胞内容物
冻融裂解[21~23]		
许多类型的细胞可以通过快速反复冻融得到裂解	细菌细胞 组织培养细胞	使用液氮，快速反复冻融至符合实验要求为止
裂解液裂解		
含去污剂的裂解液处理组织、细胞以裂解细胞膜，使细胞内容物释放出来	组织培养细胞	将细胞直接置裂解液或样品液中，进行悬浮处理
酶裂解[24,25]		
有细胞壁的细胞可以通过酶解去除细胞壁得以温和裂解	植物细胞、细菌细胞 真菌细胞	将细胞悬浮于专一性酶的等渗溶液中

（二）更强烈的破碎方法

当破碎不容易破碎的细胞如固体组织中的细胞或有坚韧细胞壁的细胞时，需要采用更强烈的破碎方法对细胞进行破碎裂解。这些方法可使细胞完全破碎裂解，但在操作过程中应避免产热及产生泡沫。主要的方法如表 13.2 所示。

表 13.2 更强烈的破碎方法

细胞破碎方法	应用对象	操作步骤
超声波处理[26~28]		
超声波产生剪切力使细胞破碎	悬浮细胞	要进行间歇处理，减轻产生热和泡沫，样品处理应在冰浴中进行
压力杯法[24,25,29]		
(French press pressure cell)用高压使细胞穿过小孔，产生剪切力从而撕破细胞	微生物细胞，如细菌、霉菌、酵母细胞及含有细胞壁的细胞	将细胞悬浮液置于预冷的压力杯中，施加压力，然后收集挤出液
研磨法[26,30~32]		
	固体组织细胞微生物细胞	组织或细胞预先用液氮冷冻，然后置瓷研钵中研磨成粉末。 加氧化铝或砂有助于研磨
机械匀浆法[20,21,33]		
要防止细胞破碎释放出蛋白酶，引起蛋白质修饰	固体组织，如心脏、肝脏、肾脏组织和细菌悬液	先尽量将组织剪成小块，然后加有蛋白酶抑制剂的冷匀浆液进行匀浆，过滤或离心收集

二、培养细胞总蛋白质的提取

裂解液裂解方法是很多实验室常用的裂解组织培养细胞的样品制备方法。下面就裂

解液裂解方法展开讨论，以下的操作过程均在冰浴中进行。这里所涉及的裂解液成分为：8mol/L 脲，65mmol/L DTT，4% CHAPS，40mmol/L Tris。

1. 胰酶酶解后裂解

细胞培养在直径为10cm的培养皿中，待培养至80%左右密度时，以含0.02% EDTA的0.05%胰蛋白酶消化，细胞经预冷的PBS漂洗3次，离心收集在1.5ml离心管中，加450μl裂解液，反复吹打。

2. 皿上直接裂解

细胞培养在直径为10cm的培养皿中，待培养至80%左右密度时，细胞用预冷的PBS漂洗3次，加450μl裂解液，细胞刮刀收集，用移液器转移至1.5ml离心管中，反复吹打。

3. 热SDS裂解[34]

细胞培养在直径为10cm的培养皿中，待培养至80%左右密度时，细胞用预冷的PBS漂洗3次，加450μl热的SDS裂解液（每10ml中，10% SDS 300μl，β-巯基乙醇 100μl，1.5mol/L Tris·HCl，pH8.8，300μl，去离子水9.3ml），细胞刮刀收集，用移液器转移至1.5ml离心管中，反复吹打后，把离心管置入液氮中迅速冷冻。在不加热情况下，真空离心冻干样品。用原来体积的样品缓冲液重新溶解样品（100ml中，尿素 59.7g，NP-40 4g，DTT 1.54g，两性电解质 pH3～10 5.5g，去离子水 44.9ml）。

最终所有的样品用超声细胞破碎仪超声处理，超声时间为5s，间歇时间为10s，功率为100～120W，超声处理至溶液清澈无黏稠物为止，处理过程在冰浴中进行。超声处理后，4℃、25 000*g* 离心1h。取上清进行蛋白质浓度测量，按实验所需的量分装后 -80℃冰箱中保存备用。如图13.2所示，A、B、C分别为以上三种方法单独制备的样品在相同的条件下得到的2-DE图谱。从图13.2可以看出，在上样量相同的情况下，胰酶酶解后裂解的样品A比B和C的蛋白质点要少，认为在胰酶酶解破碎细胞过程中有一小部分蛋白质受到了损失；热SDS裂解制备的样品C在等电聚焦方面比样品A和B要差，认为低浓

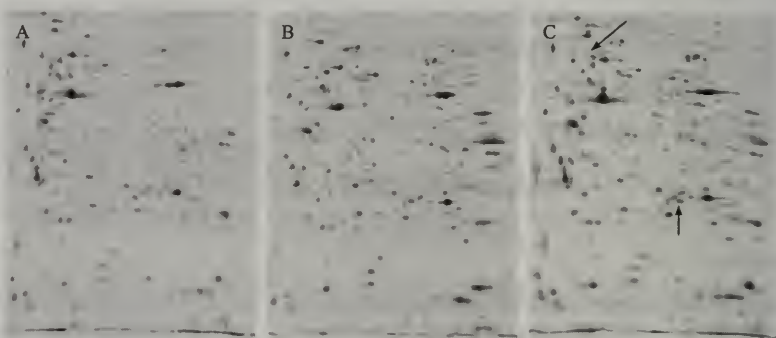


图13.2　不同的细胞裂解方法对2-DE分辨率的影响

样品取自培养的ECV 304细胞，A和B上样量均为10μg，第一向为7cm pH3－10 NL IPG胶条，第二向为11%的PAGE。A，胰酶酶解后裂解的样品的2-DE图谱；B，细胞培养皿上直接裂解的样品的2-DE图谱；C，热SDS裂解制备的样品的2-DE图谱

度的 SDS 在一定程度上同样影响了第一向等电聚焦；另外，由于 SDS 量的增大，在图谱中垂直位置上出现了一些新的点，可能是由于 SDS 包被的情况不一样所引起，如图 13.2C 中箭头所示。

以上为贴壁培养细胞的样品制备方法，至于悬浮培养的细胞（包括植物细胞），常规的处理方法就是离心收集细胞；用预冷 PBS 清洗 3 次；加裂解液进行处理；进行超声波处理；离心收集上清。注意无论是如何处理细胞，样品最终的蛋白质浓度不要低于 0.1mg/ml，最合适的浓度为 1～5mg/ml（一般来说 1×10^7个动物细胞能得到 1mg 蛋白质）。因为样品浓度过小，会使相对盐浓度增大，从而影响了等电聚焦。

第三节　组织样品制备

一、概　　述

以疾病组织为实验材料的蛋白质组学研究，因为真实地反映机体内环境的情况，既可以为疾病的早期诊断提供新的特异性的候选标志，又能够为疫苗的研制寻找新的特异性的候选抗原，同时可以监测、跟踪疾病的发展及预后情况，并为认识发病机制、明确疾病的分型、分级、分期提供线索。因此，疾病组织的蛋白质组学研究有着不可替代的临床意义及应用前景。

另外，由于水稻和拟南芥（*Arabidopsis thaliana*）这些植物基因组序列测定的完成，随之而来的问题是如何应用高通量的技术去研究其复杂的蛋白质网络。因此植物蛋白质组研究也越来越受到关注，对农业生产必将会产生巨大的影响，如现在的育种也将从以前的通过个别基因的转移来改进个别性能，发展到整体性能的改善。植物蛋白质组的研究在以农业为本的国家里，有着更丰富的内涵。然而由于动物组织与植物组织虽然同是组织样品，但是各有其物化特性，所以制备方法也不尽相同。有些制备方法可以借鉴通用，有些则必须根据其特有的性质和要求做相应的改变。下面分别介绍动物组织和植物组织样品的制备方法。

二、动物组织样品的制备

对于取材于机体组织样品的制备，不均一性和易污染性一直是技术发展的症结所在。如何从多细胞、不均一的体系中获取特定类型的细胞、亚细胞结构甚至特殊的蛋白质复合体，同时又避免各种血浆蛋白质、基质组分以及坏死组织成分的污染，是各类组织样品制备方法必须考虑的问题。早期组织样品采用酶解样品制备法（enzymatic sample preparation，ESP），即酶解、抽提细胞然后以 Percoll 梯度离心等方法分离。由于酶解处理会导致蛋白质图谱的改变，ESP 已很少使用。目前，动物组织样品常采用非酶解样品制备法、纯化法和微切割法进行制备。

（一）非酶解样品制备法（nonenzymatic sample preparation，NESP）[36]

1. 工具、设备

针头，长针头，注射器，玻璃匀浆器，解剖刀，尼龙网眼过滤器，冷冻离心机，冷冻干燥机。

2. 试剂

R/MINI1640 缓冲液（游离谷氨酰胺；5%胎牛血清；0.2mmol/L Phenylmethylsulfonyl Fluoridem；1mmol/Lol/L EDTA；抗生素：苯唑青霉素 25 μg/ml，庆大霉素 50μg/ml，青霉素 100U/ml，链霉素 100μg/ml，两性霉素 B 0.25μg/ml，Nistatin 50 U/ml），Percoll/PBS 溶液（54.7%，密度为 1.07g/ml），PBS/PIH 溶液，10% SDS/33.3% Mercapthoethanol，DNase Ⅰ溶液（0.144mg/ml 20mmol/L Tris-HCl，2mmol/L $CaCl_2 \cdot 2H_2O$，pH8.8），RNase A 溶液（0.0718mg/ml Tris），上样缓冲液（0.2mmol/L 苯甲磺酰氟 PMSF，1.0mmol/L EDTA，0.5% NP-40，25mmol/L CHAPS）

3. 操作步骤

(1) 细胞的提取

手术切除的组织块迅速置于冰上，快速切成几个肉眼可见、无坏死区域的小块（备份的原材料可用于平行的组织病理学、细胞学、免疫组化等实验）。用预冷的 R/MINI1640 缓冲液洗涤组织小块数次，根据组织病理学特征及组织块的大小，选择合适的细胞提取方法（图 13.3）。

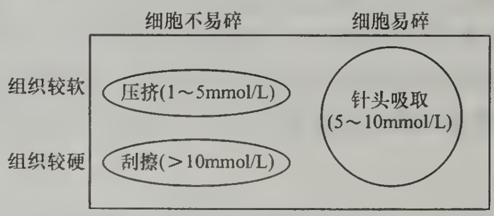


图 13.3　组织细胞提取方法

①压挤（squeezed）：在约 0.5ml 预冷的 R/MINI1640 缓冲液中，通过切、压、剁得到微小的碎片，添加缓冲液至 1～2ml。

②刮擦（scraped）：以锋利的解剖刀刀锋轻刮新鲜切出的组织小块的截面，所得细胞置于 1～2ml 预冷的 R/MINI1640 缓冲液中。

③针头吸取（Needle Aspiration）：用 0.7mmol/L 孔径的针头，收集约 1 000 万个细胞于 1～2ml 预冷的 R/MINI1640 缓冲液中。

(2) 去除组织碎片及红细胞

①先以两层直径 10mm 的尼龙网眼过滤器重悬细胞。尼龙网眼过滤器上层孔径为 250μm，下层孔径为 100μm（可根据组织细胞的实际大小选择合适的上下滤器孔径）。这一步对以刮擦法或压挤法提取的细胞悬液尤其重要，因为其中不可避免地混有黏连组织

和成纤维细胞等杂质成分。

②所得的细胞重悬液用1.2mm的长针头（连在2ml注射器上）转移至新的Effendorf管中，小心加入1.0～1.5ml预冷的Percoll/PBS溶液，使得重悬细胞浸没其中。

③于4℃，在低加减速率下（1 000g）离心10min。

(3) 收集并洗涤细胞

①用连在2ml注射器上0.7mm的针头挑取收集分界面处的细胞层，注意尽量减少Percoll及R/MINI1640缓冲液的污染。根据细胞的多少用适当体积的预冷的Percoll/PBS溶液洗涤。

②若细胞层薄甚至几乎透明，洗涤后的细胞收集在体积小于0.4ml的PBS/PIH溶液中，在盛入1.5ml预先称重的Effendorf管中，加满PBS/PIH溶液后，4℃、800g离心3min，沉淀物中的细胞颗粒重悬于PBS溶液中。若细胞层较厚，洗涤后的细胞收集在体积小于1ml的PBS/PIH溶液中，然后转移入10～12ml的玻璃离心管中，加入5～6ml预冷的PBS/PIH溶液后，4℃、800g离心3min，沉淀物中的细胞颗粒重悬于1～2个盛有预冷的PBS溶液并预先称重的Effendorf管中。

③于4℃、2 700g离心5min，小心去除所有的上清及Effendorf管内外管壁黏着的液滴。称重并计算所得细胞的净重（wet weight，WW）后－80℃冻存。

(4) 细胞裂解及核酸降解

①按细胞净重（mg）∶mQ水（μl）＝1∶1.89的比例，在冻存细胞中加入相应体积的mQ水，冰上制成细胞重悬液。反复冻融4次以破碎细胞。

注：细胞破碎还可采用超声法[37]、裂解液法[38]等。其中，裂解液法[38]用裂解缓冲液（9mol/L尿素，4％ m/V CHAP，65mmol/L DTE，4％载体两性电解质Ampholine 9—11，1mmol/L EDTA）将破细胞和溶解蛋白质两个过程结合在一起，较为简便，本实验室就是选择这种样品制备方法进行肝细胞癌的癌与癌旁组织的比较研究。

②依次加入（0.089×WW）μl 10％SDS/33.3％ Mercapthoethanol，（0.329×WW）μl DNase I溶液和RNase A溶液，冰上静置5min。

③冷冻干燥降解核酸后的样品。

④冻干的蛋白质样品溶解在（6×WW，或3×WW若WW＜10mg）μl适当的上样缓冲液中，混匀后静置3h使样品中的蛋白质充分溶解（有文献报道只需5min[35]即可）。4℃ 12 000r/min离心15min，上清小心转移入新的Effendorf管中，－80℃冻存。

（二）纯化法

由于方法所限，NESP法制备的组织样品，无法完全去除血浆蛋白质、基质组分以及坏死组织成分的污染，而且也不能排除机体组织中其他类型细胞的干扰。因此，在用NESP法制备成细胞悬液后，常常添加抗体纯化、流式细胞仪纯化或选择性的细胞培养建立不同的细胞株系。但纯化出单一类型的细胞有可能会丢失一些系统性信息，例如，细胞间基质与细胞间、不同类型细胞间的相互作用等信息。据报道，不经流式细胞仪纯化，仅用NESP法制备的组织样品中基质信息仍可以保留[40]，因此，应注意根据具体情况设计对照和并行实验。

1. 抗体纯化法(antibody purification)[38,39]

(1) 工具、设备

磁珠浓缩器(MPC-1;Deutsche Dynal),聚苯乙烯磁珠(Dynabeads M450 Pan T,Dynabeads M450 Pan B—DYNAL AS—Oslo,Norway)。

(2) 试剂

抗体、裂解缓冲液等。

(3) 操作步骤(以磁珠法为例,仅列出制备成细胞悬液后的步骤)

①依据悬液中细胞数的多少,加入适当的聚苯乙烯磁珠(带有1、2或3种抗特定细胞表面特定抗原的单抗,如抗人T细胞表面CD2膜抗原的单抗,抗人B细胞表面CD19膜抗原的单抗,抗人上皮细胞的Ber-EP4抗体)。4℃恒温30min。

②磁珠浓缩器收集挂满目标细胞的聚苯乙烯磁珠。

③洗涤并收集磁珠上的细胞,加入裂解缓冲液等进行常规制样。

注:作为对照,1.(3)(C)中以磁珠移去了目标细胞的剩余液,可以加入裂解缓冲液并进行常规制样及平行实验。

2. 流式细胞仪纯化法[40~45]

(1) 工具、设备

细胞计数器,FACS仪(如Epics,Coulter-Immol/Lunotech,Hamburg,Germany),玻璃管(如Silicoborated Glass Tube)。

(2) 试剂

台盼蓝,70%预冷乙醇,预冷PBS缓冲液,荧光素交联的抗体(如CAM5.2 Perkin-Elmer,Norwalk,Connecticut,USA)。

(3) 操作步骤(仅列出制备成细胞悬液后的步骤)

对NESP法所得的细胞悬液进行细胞计数,用台盼蓝测定细胞存活能力(一般约90%的活细胞);于4℃、70%预冷乙醇中振荡过夜;用预冷PBS缓冲液洗涤细胞;荧光素交联的抗体(抗特定细胞表面的特定抗原,如用荧光素交联的抗细胞角蛋白抗体分离结肠癌上皮细胞)标记约45min;用FACS仪定量选择被荧光素标记的细胞并收集在玻璃管中;离心、裂解细胞等常规制样操作。

(4) 备注(流式细胞术的发展)

①RCA技术与流式细胞术的联用:RCA(rolling circle amplification)[41,42]技术是基于DNA滚环复制的基本理论而设计的一种新的信号原位放大的方法,能够整合细胞及组织来源的基因组和蛋白质组信息。RCA技术有以下优点:a)线性放大(不同于PCR只能进行指数放大)原始信号(约10^5倍),因而所有的扩增产物始终连在原始目标物上,可有效消除假阳性;b)可以对目标物定位并分析生物结构;c)能够同时分析多种原始信号。目前,RCA技术已成功地应用于免疫组化、蛋白质芯片、流式细胞术等技术,大大提高了这些已有技术检测或分离的灵敏性。RCA技术与流式细胞术联用的大致步骤如下:a) RCA的准备:设计寡核苷酸序列,合成两套引物(Circles和两端联有荧光素或生物素的修饰序列);准备免疫共轭物作为RCA引物与抗体连接的中介;b)NESP法制备的细胞重

悬液以 PBS 洗涤，再连接与荧光素交联的特定单抗（如抗 CD3 的单抗 clone BMA 031；Caltag Labortories，Burlingame，CA）；c）连接免疫共轭物；d）PBS 洗涤，离心，细胞重悬于固定介质(Caltag Laboratories)，冰上固定 15min，再用 PBS 洗涤细胞；e)于 37℃、在含 200nmol/L 的 RCA 引物的 100mmol/L 谷氨酸钾溶液中温育 30min；f）RCA 反应：31℃，Φ29 DNA 复制酶催化反应 45min；g)Coulter EPICS/Profile II 流式细胞仪检测。

②激光扫描细胞术（laser scanning cytometry，LSC）[43,44]：这是一种基于显微镜可检测多种荧光嗜铬染料标记的细胞或组织样品的固相细胞术。既能极小化样品的用量，又能将荧光信息与形态学信息整合起来，因而拥有流式细胞术及图像细胞术（image cytometry）两者的特征，应用前景极为广阔。

③复合颗粒（multiplexed particle-based）流式细胞术[45]：以复合微球颗粒（已有商业化产品，如 luminex flowmetrix system）作为常规的免疫分析、亲和性分析、DNA 杂交分析的固相支持物，快速（两小时内）、无需洗涤，可在极少量的同一样品上同时得到多种不同的分析结果。

3. 建立细胞株系法

NESP 法制备成的细胞悬液经选择性的细胞培养建立不同的细胞株系，再进行破细胞、溶蛋白质和去核酸等常规制样步骤。由于经历了体外培养过程，其蛋白质图谱与直接组织来源样品的图谱有一定差别。具体方法请参阅相关文献。

（三）微切割法

1. 人工微切割法（manual microdissection）[46~50]

(1) 工具、设备

注射器，针头（规格 18～27，依实际组织细胞类型而定），Pasteur 移液管（尖端极锋利，因已由火焰烧灼并小心拉伸。另外，拉伸后的尖端黏有少许 568g/L piccolyte、437.5g/L xylene 用以黏着靶组织碎片），pipetman tip(P20)，解剖刀，光学显微镜，低温维持器。真空干燥器。

注：人工微切割工具可直接手控，或连接微机械手臂（micromanipulator arm）。

(2) 试剂

10%甲醛水溶液（或 80%乙醇、3.7%甲醛），石蜡，O.C.T 复合物，二甲苯，100%乙醇。

(3) 操作步骤

①组织切片的获得：10%甲醛水溶液（或 80%乙醇、3.7%甲醛）固定，石蜡包埋，切片（同一组织区域同时切出 5μm、10～25μm 厚的切片）。根据组织病理学特征，选择合适的固定、包埋、切片方法。例如，乳腺癌组织可用 O.C.T 复合物包埋、液氮冰冻后进行切片。

②光学显微镜下微切割：5μm 厚的组织切片（常规的 H&E Hematoxylin and Eosin 染色）置于低温维持器玻璃平面上的，先以针头（或 Pasteur 移液管，或 Pipetman Tip(P20)，或解剖刀）切去希望获得的组织部分。其余 10～25μm 厚的未染色切片整齐叠放在这片已镂空的 5μm 厚 H&E 染色切片下方，再用微切割工具获取相应部分的组织微屑。

③对人工微切割三种工具(针头、Pasteur 移液管、Pipetman Tip)的比较[48]:针头和 Pasteur 移液管都存在太锋利而不能有效地将刮取的组织大小控制在 1~2mm^2(含 500~2 000 个细胞)和切下的组织碎屑难于收集两个缺点,另外,Pasteur 移液管的玻璃尖端极易折断。Pipetman Tip(P20)既可以有效地将刮取的组织大小控制在 1~2mm^2,又可以通过存放在尖端的 0.75μl 左右的矿物油收集切下的组织碎屑(组织碎屑-矿物油混合物可反复获得并汇合在一起)。

④后处理:室温下,以二甲苯洗涤两次(每次 800μl),除去石蜡;100% 乙醇洗涤三次(每次 800μl);真空干燥去除乙醇(估计净重);加适当体积的裂解缓冲液或破细胞后加上样缓冲液;加核酸酶去 DNA 及 RNA;低温高速离心去未溶杂质。此步骤亦可参考 NESP 方法。

2. 激光微切割(laser microdissection)[51~58]

激光即 light amplification by stimulated emission of radiaton (Laser),按介质大致分为四类:固体激光,如红宝石激光;液体激光,如染料激光;气体激光,如二氧化碳激光、氩激光;半导体激光,如二极体激光。1964 年,Tomberg 首先将单脉冲激光用于诱导细胞损伤,1970 年 Bessis、1971 年 Hillenkamp、1972 年 Bereiter-Hahn 等人也用红宝石激光(ruby laser)进行了尝试。但真正意义的激光微切割仪器是 Remy 等人在 1975 年发展的,使用氮激光(nitrogen laser)取代了破坏性太强的红宝石激光。1976 年 Meier-Ruge 发展了紫外激光仪,发射低损伤的 0.5mW 氦氖激光束(He-Ne laser)来进行微切割。随后 20 多年,各种激光和相应的仪器蓬勃发展起来。从方法学上讲,激光微切割技术可分为正选择靶点、负剥离非目标区域两大类。其中,最具代表性的是两种正选择技术:基于红外激光(infrared laser)的激光捕获显微解剖技术(laser-capture micro dissection, LCM)和基于氮激光(nitrogen laser)的激光微束微切割-激光压力弹送法(laser microbeam microdissection-laser pressure catapulting, LMM-LPC)。

(1) 激光捕获显微解剖技术(laser-capture microdissection, LCM)[51~56]

①工具、设备

透明热敏膜(transparent thermoplastic film, 100μm 厚,材质为 EVA ethylene vinyl acetate polymer), Pixcell™ Laser capture microdissection system (Arcturus Engineering 公司产品)

②操作步骤

a) 常规组织切片或冰冻切片及对照用同源组织染色切片按如下方法准备:甲醛或乙醇固定、石蜡包埋的 6~8μm 厚的常规组织切片;新鲜组织块液氮冰冻,低温制成 6~8μm 厚的冰冻切片;对照用同源组织切片脱石蜡、H&E 染色、3% 甘油处理 1min 并风干 5min。上述所有切片梯度酒精脱水,风干 5min。

b) 要进行 LCM 的常规组织切片或冰冻切片覆盖一层热敏膜,制成"膜-组织-玻片"三明治(图 13.4);Pixcell™ laser capture microdissection system 中选择内置显微镜的合适放大倍数观测"膜-组织-玻片"三明治。

c) 通过与冷凝器光路共轴的一块前表面反射镜,在热敏膜上表面输入脉冲 CO_2 激光束;CO_2 激光(Apollo Company Model 580, Los Angeles 或 California Laser Company Model

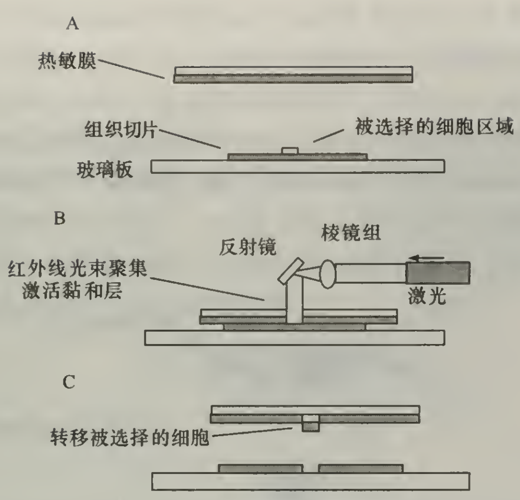


图 13.4 “膜-组织-玻片”三明治

LS 150, San Marcos, CA)的脉冲波长及功率均可以根据情况调节。

d) ZnSe 棱镜将 CO_2激光束聚焦在组织切片的目标点上，产生的照射点直径可调。例如：600ms、25～30mW 的脉冲激光束可产生直径 150μm 的照射点。

e) 激光束 90%以上的能量被 100μm 热敏膜吸收（10.6μm 波长时，热敏膜的吸收指数为 $200cm^{-1}$，仅少量热能被膜下样品吸收；因而，膜下样品可以耐受脉冲激光束产生的瞬间 90℃的高温而无物质变化)，样品下玻片的存在也可确保热量沉积于靶点处整个 100μm 厚的热敏膜层，导致膜融化，润湿靶点处目标样品，迅速冷却和再结晶。

f) 在靶点处，膜与目标样品的黏着力强于目标样品与玻片间的黏着力，可通过机械系统转移并收集膜与目标样品。一般而言，每 1 000～3 000 束照射收集满一个塑料帽(30min 左右)，每次激光照射平均可捕获 3～5 个细胞，每次免疫杂交需要 20 000～40 000 个细胞，耗时 2～7h；每次双向电泳需要约 250 000 个细胞，耗时约 15h。

③备注

LCM 技术的主要优势在于：a)直接的可视显微环境下，无污染而一步获取目标物；b)广泛适用于冰冻组织切片、石蜡包埋组织切片、甲醛或乙醇固定后经染色(H&E 等方法)或未经染色的组织切片；c)当前，脉冲激光束可捕获直径 30～700μm 范围内的靶细胞，由于激光束可以聚集直径小到 10μm 的单细胞，因而，不久的将来 LCM 可以实现更为精细的显微切割。

LCM 技术的缺点：a)耗时；瓶颈在于当前蛋白质组学支撑技术之一的双向电泳技术的检测极限是 1ng(银染)，这要求蛋白质丰度范围在每细胞 50 000～100 000 个拷贝，细胞内低拷贝蛋白质因为这个技术缺陷难于检测，也因此对制备样品的细胞数目有依赖性；b) 脉冲激光束产生直径 30～700μm 的圆形照射区域，对组织细胞的形状有所限制，如梭形、多角形等不规则形状的组织细胞；c)红外波长激光的产热问题，虽然由于 100μm 热敏膜的存在而缓解，但不能完全保证 90℃的瞬间高温对样品毫无影响。

本实验室已与上海东方肝胆医院合作开展，利用 LCM 技术研究肝细胞癌的工作。

(2) 激光微束微切割-激光压力弹送法(laser microbeam micro dissection-laser pressure catapulting, LMM-LPC)[57,58]

①工具、设备

载玻片，Microbeam-MOMeNT 技术，1.35μm 厚的聚乙烯箔(PALM, Wolfratshausen, Genmany)，Robot-MicroBeam (PALM)显微切割仪，微离心管帽或玻璃片(联在 LPC 收集器上，P. A. L. M. GmbH)。

②操作程序(图 13.5)

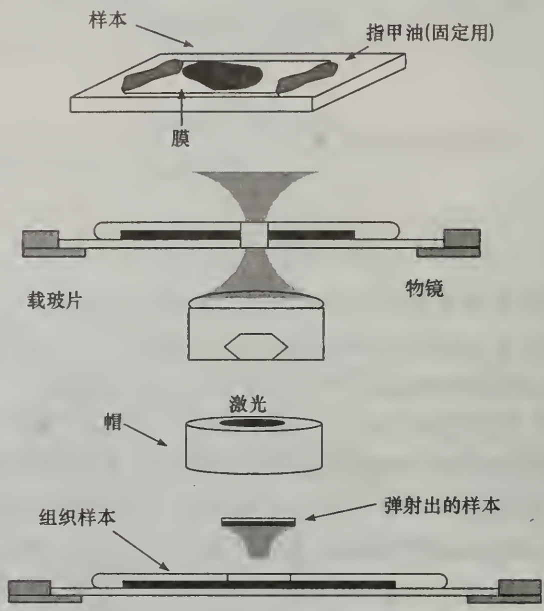


图 13.5 LPC 操作程序示意图

a) 组织切片的准备：常规甲醛固定、石蜡包埋的组织切片直接放置在玻璃载玻片上或间隔 1.35μm 厚的聚乙烯膜(PALM, Wolfratshausen, Genmany)；组织切片进行脱石蜡和 H&E 等染色处理；建议在脱石蜡和染色处理后，将聚乙烯膜和组织切片的上下位置交换，即聚乙烯膜覆盖在组织切片上，这样既可以减少污染，又可以作为微切割之后的 LPC 的组织支持骨架以保证所获细胞的形态学完整性；聚乙烯支持膜的安放可使用 Microbeam-MOMeNT 技术，每块膜可在 1min 内完成；聚乙烯膜覆盖的未脱石蜡和染色处理的组织切片可保存在常温。

b) LMM 和 LPC：选择合适的物镜，并根据组织特性及物镜传输效率调节所需激光束的能量；对 LMM 而言，激光束能量需要调整到仅切割组织片或聚乙烯膜－组织片(根据组织切片是否是直接放置在载玻片上)，应视组织与激光束具体相互作用情况而定；对 LPC 而言，激光束能量需要调整到能够弹送并收集所选的组织小片到微离心管帽或玻璃片(联在 LPC 收集器上，P. A. L. M. GmbH)上。一般来说，LPC 激光束能量是 LMM 激光束能量的两倍；使用的 Robot-MicroBeam (PALM)显微切割仪可以发射能量可调的 337nm 脉冲氮激光束并含内置显微镜(Axiovert 135; Carl Zeiss, Oberkochen, Germany)。

c) 最终收集的所有组织小片进行常规的裂细胞、溶蛋白质、去核酸等处理。

③备注

LMM-LPC 技术的主要优点在于：a)与待切割组织无直接的机械接触以保持组织的原始形态；b)微切割选择的部分与未被选择的部分之间有 1～10μm 的间隙可避免污染并保持组织的原始形态；c)在微切割的同时用 LPC 技术将所获细胞或细胞簇传送到容器内；d)对组织细胞形状和大小无限制，梭形、多角形等不规则形状的组织细胞及直径从一到数百微米的组织细胞都可以获取；e)获取及传送速度较快，单细胞 30s 内完成，直径 100μm 内的细胞簇在 60s 内完成，直径 1 000μm 的细胞簇需 1～2min；f)337nm 氮激光不需加热就可以严格定位并充分光解待剥离部分。

（四）其他

1. 组织芯片(tissue chip)[59～61]

(1) 工具、设备

钢制活塞钻孔器(薄壁钢管围成的内径为 600μm，自带钢探针)，受体石蜡块(点阵状排列、深度可调的内径为 600μm 的孔穴)，黏性带系统(Instrumedics, Hacken-sack, New Jersey)。

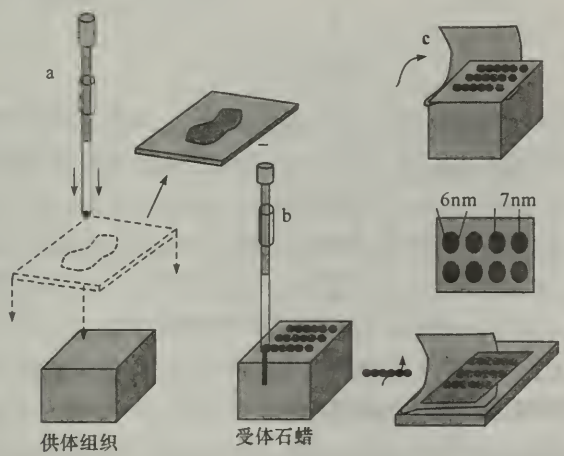


图 13.6　组织芯片操作步骤

(2) 操作步骤[59](图 13.6)

①不同来源的新鲜组织块用乙醇固定或甲醛固定。

②准备各组织块打洞截面的 H&E 染色切片作为打洞获取圆柱状组织核时的可视参照。

③用钢制活塞钻孔器的薄壁钢管(无污染、锋利)在供体组织块的目标部位打洞以截取圆柱状组织核。

④钢探针将截取的圆柱状组织核压出薄壁钢管，并按照其 X-Y 轴坐标准确定位于受体石蜡块的对应孔穴内。可以使用立体测量显微镜(stereotactic microscope)来保证组织核放置的高度有序性。

⑤清洗钢制活塞钻孔器，重复③④操作，直至受体石蜡块的孔穴内排满组织核(来源

于相同供体组织块和(或)不同供体组织块)。一般地,每块表面积 45mm×20mm 的受体石蜡块上可排列 1 000 个组织核(组织核间距 100μm)。

⑥所得的组织核微阵列块用超薄切片机切出一系列 8μm 组织阵列片,然后进行常规的裂解细胞、溶蛋白质、去核酸等处理。

注意事项 ① 切片时建议使用黏性带系统(Instrumedics, Hackensack ,New Jersey),可以避免组织核移位;即:黏性带覆盖在组织核微阵列块上,使二者相互黏着,再切割紧贴黏性带的下方,所得的组织阵列片粘于黏性带表面,随后转移组织阵列片并移去黏性带。② 平均每块组织核微阵列块可切出 200 个组织阵列片;H&E 染色能够保证组织学图谱同源性的厚度依组织类型有所差别:对迁移后癌组织而言,可保证 3mmol/L 厚(约 40 个 8μm 组织阵列片);对于高分化癌组织、轻微损伤组织及正常组织而言,可保证 1～2mmol/L 厚(约 20 个 8μm 组织阵列片)。

(3) 备注

激光微切割等技术可以和组织芯片技术联用,例如 LCM 与组织芯片的联用[60](如图 13.7)。

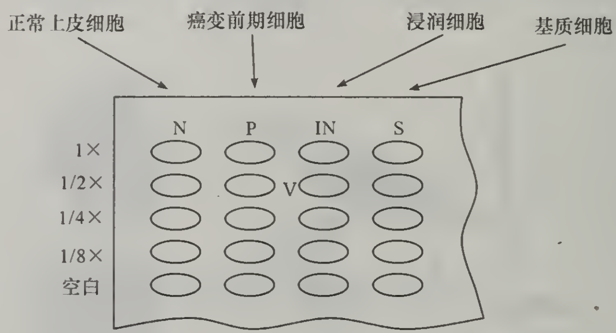


图 13.7　LCM 与组织芯片的联用

2. 质谱成像术(imaging mass spectrometry)[62～64]

基质辅助激光解析电离-飞行时间质谱(MALDI-TOF MS)的一大优势在于能够在纯度不高的样品中分析相对复杂的复合物;电杂交、接触杂交等技术可以将蛋白质样品转移到 poly-vinylidene fluoride(PVDF)、Nylon、polyethylene 或 Teflon 膜上,然后直接进行质谱分析。基于这两方面的技术,直接大规模地分析生物样品成为可能,使得能够分析单细胞、细胞簇、新鲜组织或常规组织切片并获取平面或空间信息的质谱成像术应运而生。同其他的成像技术相比,质谱成像术的成像不局限在特异的一种或几种蛋白质分子,理论上可以找到组织切片中的每个已知或未知的蛋白质分子。

(1) 工具、设备

MALDI-MS(Voyager Elite DE, Applied Biosystems, Framingham, Massachusetts),载玻片,聚乙烯(Polyethylene)膜,Perseptive DE-STR MALDI-MS。

(2) 试剂

自动吸量管,基质缓冲液(Sinapinic Acid: 10mg/ml Acetonitrile/H_2O = 1∶1, V/V),胰蛋白酶溶液(终浓度 2 pmol/μl 溶于 50mmol/L ammonium bicabonate),mQ 水。

(3) 操作步骤

①以冰冻组织切片为例[64](图 13.8)

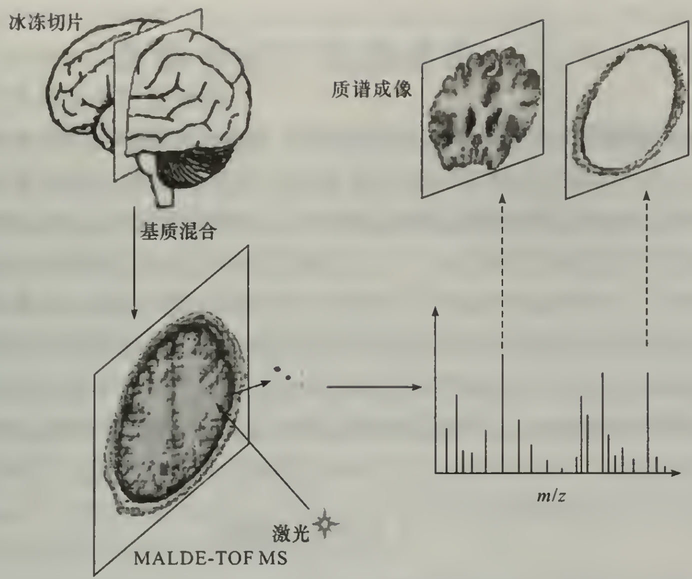


图 13.8 质谱成像术

a) 将于液氮保存的新鲜组织,切割成连续的 12μm 厚冰冻组织切片。

b) 无污染的目标钢板上放置一片冰冻组织切片,覆盖基质缓冲液后脱水。

c) MALDI-MS(Voyager Elite DE ,Applied Biosystems,Framingham, Massachusetts)分析:样品表面的每个试映图(raster)以连续的激光点(直径 25μm)照射,固定激光位置,移动样品盘;每个激光照射点产生照射部分的样品信息。平均每 50 次激光照射产生一个 fast transient recorder PC board (DP211, Acqiris,Geeva, Switzerland),每个照射点信息的处理需要约 2.5s;一个典型的信息点阵根据成像质量要求需要 1 000～30 000 个照射点;每个照射点可产生 500～80 000Da 的约 200 种蛋白质或肽段峰。

d) 软件分析,生成样品的质谱图像或任意特定相对分子质量蛋白质或肽段的样品定位图。

②以新鲜组织为例(建议在同一天内完成下列操作)[63]

a) 于密闭 CO_2环境麻醉实验动物,快速活体解剖取出目标组织并切割成合适大小的组织片。

b) 组织片置于载玻片上,其新鲜截面平整覆盖上一层聚乙烯(Polyethylene)膜,最大限度减少水分的丢失。

c) 接触杂交 5min 后移去组织片,揭下聚乙烯膜并使之完全失水。

d) 每一个杂交区域用 20μl mQ 水润洗,20μl automatic pipet 除去盐和组织碎屑,脱水。

e) 1μl 胰蛋白酶溶液(终浓度 2 pmol/μl 溶于 50mmol/L ammol/Lonium Bicabonate)点在杂交区域上,37℃ 酶解 30min,其间定时滴加 mQ 水以保持湿润。

f) 聚乙烯膜重新失水,在杂交区域点上 1μl 基质(Sinapinic Acid: 10mg/ml acetoni-

trile/H_2O=1∶1, V/V),环境空气中晾干。

g) 用 Perseptive DE-STR MALDI-MS 分析样品。

三、植物组织样品的制备

上面提到的动物组织样品的制备方法有些对植物组织同样适用,但应根据植物组织的特点,在应用时而做相应的改进。由于植物组织含有坚韧的细胞壁和复杂的细胞外基质等,所以,在样品制备时要采用更强烈的破碎方法才能破碎细胞获得样品。对于植物材料来说,标准的蛋白质提取方案是在 TCA/丙酮溶液中沉淀蛋白质,然后用含去污剂的裂解液来溶解沉淀。这种方法的优点是:克服在单位质量的植物组织中蛋白质提取量低的问题;去除了植物次级代谢产物带来的影响;使植物体内高活性的蛋白酶失活,避免了蛋白质被降解修饰。其缺点在于由于不充分地沉淀和溶解而选择性地丢失某些蛋白质。最近 Giavalisco[65]等改善了 Klose[66]新近发表的对动物组织抽提的方法,使之适用于植物组织的蛋白质分级提取。下面分别介绍这两种从植物组织中抽提蛋白质样品的方法。

(一) TCA/丙酮法

简单的植物样品制备可在加裂解液的情况下,在 TCA 的丙酮溶液中进行匀浆破碎处理,然后离心获得。下面以拟南芥幼叶(green leave)为例介绍,具体操作步骤[67]。

①称取新鲜的嫩叶 200mg,洗净。

②使用液氮速冻,进行研磨,加含 0.07%巯基乙醇,10%TCA 的丙酮溶液,低温研磨 45min。

③35 000g 离心 15min,用 0.07%巯基乙醇,1mmol/L PMSF 和 2mmol/L EDTA 溶液清洗沉淀,然后冻干。

④每 10mg 冻干粉末中加 250μl 样品裂解液(9mol/L 尿素,2%两性电解质 pH3~10,60mmol/L DTT,0.5% Triton X-100 和 0.003%溴酚蓝)。

⑤搅拌情况下在 37℃温浴 30min。

⑥19 000g 离心 15min,取上清。

(二) 植物组织分级提取法

同样以拟南芥幼叶为例[65],介绍一种从 2-DE/MS 路线出发的植物组织分级抽提的样品制备方法。幼叶取自从胚芽算起生长了 56d 的植株,双蒸水清洗后用滤纸去掉多余的水分。在进行蛋白质抽提前,幼叶用液氮速冻保存在 -80℃冰箱中。蛋白酶抑制剂复合物分为三种:复合物Ⅰ为一片 Complete™(Molecular Biochemicals Roche, Mannheim, Germany)溶解在 2ml 含 100mmol/L KCl,20%甘油,50mmol/L Tris,pH7.1 的溶液中;复合物Ⅱ为 1mmol/L 胃酶抑制剂 A 和 1.4μm PMSF 的乙醇溶液;复合物Ⅲ为一片 Complete™溶解在 2ml 含 0.2mol/LKCl,20%甘油,0.1mol/L 磷酸,pH7.1 的溶液中。具体的操作过程如下所述。

①取冻存的植物组织 20~300mg,加 1/8 重量的抑制剂复合物Ⅰ和 1/20 重量的抑制剂复合物Ⅱ。

②液氮浴中匀浆成碎末。

③在 4℃下 226 000g 离心 30min。

④上清即为分离所得的第一个组分，每 50μl 加入 54mg 脲，5μl 700mmol/L DTT，5μl 两性电解质复合物 Servalyte 2－4。如果不是立即进行等电聚焦，样品在－80℃保存。所获得的蛋白质主要是疏水性蛋白质。

⑤用 1/8 重量的抑制剂复合物Ⅲ，相同重量的 100mmol/L 磷酸缓冲液含 0.2mol/L KCl，20％甘油，4％ CHAPS 和 2％的 ASB 14 去溶解沉淀。

⑥重复步骤 2，在 4℃下搅拌 45min，加 23％体积质量比的 700mmol/L DTT 溶液，7mol/L 脲和 2mol/L 硫脲。

⑦常温下搅拌 45min。

⑧当脲和硫脲完全溶解时，在 17℃下 226 000g 离心 30min。

⑨上清即为分离所得的第二个组分，加 0.05 体积质量比的两性电解质 pH2－4，样品可以直接进行等电聚焦，或者在－80℃保存；获得的蛋白质主要是疏水性蛋白和核酸结合蛋白。

⑩加含 2mmol/L$MgSO_4$的 50mmol/LTris pH8.0 溶液溶解沉淀，重复步骤 2。

⑪加 0.025 体积质量比的 DNase 搅拌 30min。

⑫加 0.11 体积质量比的 700mmol/L DTT 溶液，7mol/L 脲和 2mol/L 硫脲，搅拌 30min。

⑬当脲和硫脲完全溶解时，在 17℃下 226 000g 离心 30min。

⑭上清即为分离所得的第三个组分，加 0.05 体积质量比的两性电解质 pH2－4，样品可以直接进行等电聚焦，或者在－80℃保存；获得的蛋白质主要是与核酸结合紧密的蛋白质。

由于植物组织纤维素含量较高，一般取植物纤维素含量低的部位制备样品，如嫩叶、幼根、种子等。另外，植物的亚细胞和复合物分级研究也开始兴起，如细胞膜蛋白、线粒体、叶绿体和磷酸化蛋白质复合物等，具体方法与动物组织的亚细胞和复合物分级基本上相同。我们可以借鉴动物组织样品制备方法，对不同的植物组织采用最适宜的制备方法。

四、注意事项

一般而言，无论动物组织还是植物组织，在样品制备时都必须注意以下几点。①尽量使用新鲜的组织制样，如需保存，则应该迅速冷冻保存在液氮或－80℃低温冰箱中。②使用预冷的缓冲液，整个操作过程应该在冷库或者冰浴中进行，以避免随机的物质降解。③大部分的非目标组织应先切成小块并洗涤除去。④使用蛋白酶抑制剂，避免蛋白质的降解和修饰。⑤少量的核酸不会引起蛋白质聚集影响等电聚焦，但是大量的样品制备时必须用核酸酶进行处理；另外也要去除多糖等干扰分子。

第四节　分泌蛋白质样品的制备

细胞分泌蛋白质主要包括细胞因子、生长因子和激素等具有重要功能的生物活性分

子。分泌蛋白质的重要性决不亚于胞内蛋白质或膜上受体蛋白质，如致病细菌的分泌蛋白质往往含有致病或毒性因子，对于疾病的病理研究和抗生素开发[68]等具有重要意义；真核细胞的分泌蛋白质负责细胞的自反馈调节以及细胞间的近程或远程的信号传导，在研究信号传导通路、癌细胞生长转移和免疫调节等方面具有十分重要的作用。因此，借鉴蛋白质组的重要意义和手段，分泌组（secretome）这一概念也应运而生[69]，即指一个有机体所有的分泌蛋白质。

一、真核细胞的分泌蛋白质样品制备

培养细胞的分泌蛋白质就在细胞培养液中，收集分泌蛋白质要注意两点：一是细胞需无血清培养基培养，或在血清培养基中培养至一定密度后无血清培养，以避免血清中的蛋白质混在分泌蛋白质中；二是脱盐和浓缩，由于培养液中有大量的盐，并且体积较大，分泌蛋白质相对浓度很低。通常采用分子超滤膜，将脱盐和浓缩一次完成，具体操作步骤如下所述。

①细胞在含血清的培养基中培养（或直接用无血清培养基培养）。

②待细胞密度培养至 80%左右时，换无血清培养基，处理 6h。

③收集培养液，离心去除细胞碎片，取上清，加蛋白酶抑制剂。

④使用 Millipore 公司的 15ml 规格的 YM-3 超滤管进行超滤浓缩，降低相对盐浓度。

⑤浓缩样品经稀释后，再超滤，降低绝对盐浓度。

⑥最终浓缩样品加常规裂解液，蛋白质定量后，分装冻存。

一般而言，200ml 培养液中只能提取到 2mg 左右的蛋白质。对体积如此之大的培养液进行超滤需要很长时间。为了防止蛋白质降解和修饰，离心超滤过程应该在 4℃进行。图 13.9 为脂肪细胞分泌蛋白质的 2-DE 图谱。

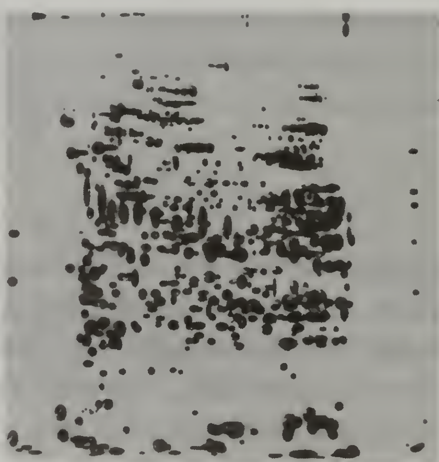


图 13.9　脂肪细胞分泌蛋白质图谱

第一向为 13cm pH 3～10 NL IPG 胶条，第二向为 12.5%的 PAGE

二、细菌分泌蛋白质样品制备

细菌蛋白质分泌是一个复杂的多阶段的过程，对细菌的膜和细胞壁的生物合成以及细菌的存活是必不可少的。Economou[70]把细菌分泌途径分成三个明显的阶段，第一阶段为靶向阶段，在导航因子的帮助下，靶向细胞膜；第二阶段为穿膜阶段；第三阶段为成熟释放阶段。一般意义上获得的分泌蛋白质为成熟释放到培养液中的分泌蛋白质。下面以幽门螺杆菌分泌蛋白质样品的制备为例[71]，介绍细菌分泌蛋白质样品的制备过程。

①细菌培养到一定密度后，在 4℃下 20 000g 离心 15min。

②取上清，0.45μm 滤膜过滤，去除残余的细菌。

③每 380ml 滤液加 120ml 预冷的浓度为 25%的 TCA 溶液沉淀蛋白质，冰浴 15min。

④在 4℃下 10 000g 离心 10min。

⑤用 10ml 丙酮重悬沉淀，清洗 2 次。

⑥冻干沉淀。

分泌蛋白质样品的制备，主要在于浓缩蛋白质和除盐。由于在培养液中分泌蛋白质的量相对很少，因此在提取分泌蛋白质时，一要注意提取分泌蛋白质前要使用无血清的培养液，避免血清中成分对分泌蛋白质的覆盖；二要注意把细胞扩增至一定的量后再进行分泌蛋白质提取，或者进行多次提取，确保获得的分泌蛋白质的量达到实验所需。至于除盐和浓缩的方法，可以根据自己的实验要求进行调整，为后续的分离和鉴定方法做好准备。

第五节　体液蛋白质样品的制备

体液是机体内所有物质交换和运输的场所，包括细胞内液和细胞外液。一般意义上的体液是指细胞外液，包括血液、脑脊液、尿液、唾液、泪水、乳汁、精液、胆汁等。体液蛋白质分析对于疾病的诊断、疾病的机制研究和疾病治疗、疫苗和药物的开发等都有着重要的意义，借助于有着高通量分析能力的蛋白质组研究手段和方法，体液蛋白质组学研究正在蓬勃发展。下面以血清和脑脊液为例，介绍体液蛋白质样品的制备。

一、血清蛋白质样品的制备

血清是成分十分复杂的样品，除了含有丰富的蛋白质（每毫升血清中含有 60～80mg 的蛋白质）外，还含有很多小分子物质，如盐类、脂类、氨基酸和糖类等。其中的蛋白质大部分是白蛋白、免疫球蛋白、转铁蛋白、结合珠蛋白和脂蛋白等少数种类的蛋白质，如果不加处理，将严重影响对种类上占绝大多数的低丰度蛋白质的检测。然而由于血清中极高丰度的蛋白质浓度与低丰度的蛋白质浓度的比值可达 10^9 倍，因此单独去除高丰度蛋白质并不能达到很好的效果，还需要对样品进行分级处理，把样品分成尽可能多的组分。

Joshua[74]等人先将血清样品过蛋白质 A/G 的亲和柱，除去免疫球蛋白这些高丰度蛋白质，再经胰蛋白酶酶解后用强阳离子交换色谱将样品分成 120 个组分，通过这些分离手段和后续的鉴定方法，共检测了 490 个蛋白质。在去除血清中高丰度蛋白质方面，许多公

司都有操作方便的试剂盒,下面以 Millipore 公司的 Montage™ Albumin Deplete Kit 为例作一介绍。图 13.10 是这个试剂盒的简易操作过程,图 13.11 显示应用此试剂盒,提高了 2-DE 的分辨率,发现了更多的低丰度蛋白质。因此,我们可以应用这些试剂盒来去除白蛋白,随后样品再用 8mol/L 脲, 65mmol/L DTT, 4% CHAPS, 40mmol/L Tris 溶解,进行常规双向电泳。

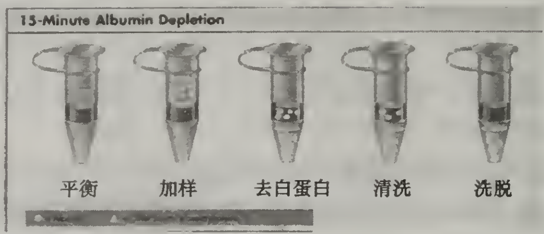


图 13.10 Montage™ Albumin Deplete Kit
的简易工作原理示意图

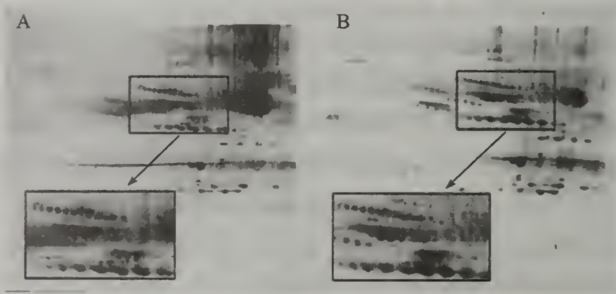


图 13.11 A 为使用 Montage™ Albumin Deplete Kit 前的 2-DE 图谱;
B 为使用 Montage™ Albumin Deplete Kit 后的 2-DE 图谱

二、脑脊液蛋白质样品的制备

脑脊液中富含多种盐类对相对少量蛋白质的分离有很大的影响。Yuan[75]等人的工作很有特色,他们在采用常规的 2-DE-MALDI-MS 路线之前,加入了三种不同的除盐方法制备脑脊液蛋白质样品。下面分别介绍这三种除盐方法。

(一)丙酮沉淀法

①100μl 脑脊液在终浓度为 80% 的丙酮溶液中 −20℃ 过夜。
②蛋白质沉淀用 80μl 浓度为 90% 的丙酮重悬,清洗 2 次。
③蛋白质沉淀用 10μl 含 5%SDS 和 2%DTT 的溶液溶解。
④于 95℃ 加热 5min。
样品与重泡胀液混合后,SDS 的最终浓度为 0.14%,不会影响等电聚焦。

（二）TCA/丙酮沉淀法

①100μl 脑脊液在 TCA/丙酮溶液中 4℃处理 1h，TCA 的终浓度为 20%，丙酮的终浓度为 80%。

②蛋白质沉淀用 80μl 丙酮重悬，清洗 2 次。

③蛋白质沉淀用 10μl 含 5%SDS 和 2%DTT 的溶液溶解。

④于 95℃加热 5min。

（三）Bio-Spin 柱除盐法

Bio-Spin 柱为 Bio-Rad 公司的产品，每 100μl 脑脊液使用分子质量 6 000Da cut-off 的微型柱，在 1 000*g* 下离心 4min。分子质量小于 6 000 Da 的蛋白质、盐类和其他干扰物保留在柱子中，分子质量大于 6 000 Da 的蛋白质被收集下来，制成样品。

Yuan 等通过对上面三种方法的比较，发现 Bio-Spin 柱预处理法蛋白质损失量最低，且在 2-DE 图谱上蛋白质点最多，是一种很值得推荐的样品脱盐方式。另外，要获得更好的分辨率，对脑脊液样品也可以进行分级处理等。

其他体液样品可用上述类似的方法进行处理，主要包括三方面内容：一是去除高丰度蛋白质，提高对低丰度蛋白质的检测；二是降低体液样品中的盐浓度，提高等电聚焦效果；三是尽可能分成多个组分，富集低丰度蛋白质，提高分辨率。

三、总　　结

在“蛋白质组学”一词诞生开始，蛋白质组学研究就同时向两个方向发展。一是走“竭泽法”的路线，即采用高通量的蛋白质组研究技术分析生物体内尽可能完整的蛋白质。这就要求在样品制备中，使用溶解性最好的裂解液配方，另外，对样品进行尽可能多的组分分级，以期获得细胞中所有的蛋白质；二是走功能蛋白质组学这条路线，把蛋白质组学作为一种手段和方法来研究生命的本质。这时的样品制备方法则必须根据需要来制定，围绕所研究的对象进行。另外，由于蛋白质鉴定技术的发展如“Shotgun”质谱检测技术、MALDI-TOF-MS 和微流芯片等，出现了一些非常规的样品制备方法。可见，蛋白质组样品制备在某种意义上来说是一种艺术，其中融合了各种各样样品的物化特性，各不相同的实验需求和不同的研究手段和方法等。在蛋白质组研究蓬勃发展的今天，样品制备或者说样品的预处理已经成为蛋白质组学研究的一个瓶颈，对研究者提出了更高的要求。

（屠成剑　李　辰）

参考文献

1 Tang W, Harrata AK, Lee CS. Two-dimensional analysis of recombinant *E. coli* proteins using capillary isoelectric focusing electrospray ionization mass spectrometry. Anal Chem, 1997, 69: 3177～3182

2 Link AJ, Eng J, Schieltz DM *et al*. Direct analysis of protein complexes using mass spectrometry. Nat Biotech, 1999, 17: 676～682

3 Pinto DM, Ning Y, Figeys D. An enhanced microfluidic chip coupled to an electrospray Qstar mass spectrometer for pro-

tein identification. Electrophoresis, 2000, 21:181～190
4 O'Farrell PH. High resolution two-dimensional electrophoresis of proteins. J Biol Chem, 1975, 250: 4007～4021
5 Molloy MP, Herbert BR, Walsh BJ *et al*. Extraction of membrane proteins by differential solubilization for separation using two-dimensional gel electrophoresis. Electrophoresis, 1998,19:837～844
6 Klose J, Kobalz U. Two-dimensional electrophoresis of proteins: an undated protocol and implications for a functional analysis of the genome. Electrophoresis, 1995, 16: 1034～1059
7 Rabilloud, T. Solubilization of proteins in 2-D electrophoresis: an outline. Methods Mol Biol, 1999,112:9～19
8 Macri J, McGee B, Thomas JN *et al*. Cardiac sarcoplasmic reticulum and sarcolemmal proteins separated by two-dimensional electrophoresis: surfactant effects on membrane solubilization. Electrophoresis, 2000,21:1685～1693
9 Molloy MP. Two-dimensional electrophoresis of membrane proteins using immobilized pH gradients. Anal Biochem, 2000,280:1～10
10 Link AJ. Internal Standards for 2-D Methods Mol Biol,1999,112:281～284
11 Pennington SR, Dunn MJ. Proteomics. From protein sequence to function. New York: Springer/Bios,2001
12 Rabilloud T. Proteome research: two dimensional gel electrophoresis and identification tools. Berlin:Springer,2000
13 Pasquali C, Fialka I, Huber LA. Preparative two-dimensional gel electrophoresis of membrane proteins. Electrophoresis, 1997, 18: 2573～2581
14 Rabilloud T, Adessi C, Giraudel A *et al*. Improvement of the solubilization of proteins in two-dimensional electrophoresis with immobilized pH gradients. Electrophoresis, 1997, 18: 307～316
15 Görg A,Obermaier C, Boguth G *et al*. The current state of two-dimensional electrophoresis with immobilized pH gradients. Electrophoresis,2000, 21:1037～1053
16 Herbert BR *et al*. Improved protein solubility in two-dimensional electrophoresis using tributylphosphine as reducing agent. Electrophoresis, 1998, 19:845～851
17 Harder A, Wildgruber R, Nawrocki A, *et al*. Comparison of yeast cell protein solubilization procedures for two-dimensional electrophoresis. Electrophoresis, 1999, 20: 826～829
18 Chevallet M, *et al*. New zwitterionic detergents improve the analysis of membrane proteins by two-dimensional electrophoresis. Electrophoresis , 1998, 19:1901～1909
19 Molloy MP, Herbert BR, Williams KL *et al*. Extraction of Escherichia coli proteins with organic solvents prior to two-dimensional electrophoresis. Electrophoresis, 1999, 20: 701～704
20 Dignam JD. Preparation of extracts from higher eukaryotes. Methods Enzymol, 1990,182:194～203
21 Lenstra JA, Bloemendal H. Topography of the total protein population from cultured cells upon fractionation by chemical extractions. Eur J Biochem, 1983, 135:413～423
22 Bollag DM, Edelstein SJ. Protein methods. New York: Wiley-Liss, 1991
23 Toda T, Ishijima Y, Matsushita, H *et al*. Detection of thymopoietin-responsive proteins in nude mouse spleen cells by two-dimensional polyacrylamide gel electrophoresis and image processing. Electrophoresis , 1994, 15:984～987
24 Cull M, McHenry CS. Preparation of extracts from prokaryotes. Methods Enzymol, 1990, 182:147～153
25 Jazwinski SM. Preparation of extracts from yeast. Methods Enzymol , 1990,182: 154～174
26 Görg A, Postel W, Günther S. The current state of two-dimensional electrophoresis with immobilized pH gradients. Electrophoresis, 1988, 9:531～546
27 Kawaguchi SI, Kuramitsu S. Separation of heat-stable proteins from Thermus thermophilus HB8 by two-dimensional electrophoresis. Electrophoresis, 1995, 16: 1060～1066
28 Teixeira-Gomes AP, Cloeckaert A, Bezard G *et al*. Mapping and identification of Brucella Melitensis proteins by two-dimensional electrophoresis and microsequencing. Electrophoresis, 1997,18:156～162
29 Ames GFL, Nikaido K. Two-dimensional gel electrophoresis of membrane proteins. Biochemistry , 1976,15:616～623
30 Görg A,Boguth G, Obermaier C *et al*. Two-dimensional polyacrylamide gel electrophoresis with immobilized pH gradients in the first dimension(IPG-DALT): the state of the art and the controversy of vertical versus horizontal systems.

Electrophoresis, 1995,16:1079～1086

31 Görg A, Postel W, Domscheit A *et al*. Two-dimensional electrophoresis with immobilized pH gradients of leaf proteins from barley (Hordeum vulagre):method, reproducibility, and genetic aspects. Electrophoresis , 1988,9:681～692

32 Posch A, van den Berg BM, Burg HCJ *et al*. Genetic variability of carrot seed proteins analyzed by one- and two-dimensional electrophoresis with immobilized pH gradients. Electrophoresis,1995,16:1312～1316

33 Geigenheimer P. Preparation of extracts from plants. Methods Enzymol,1990,182: 174～193

34 Garrels JI. Quantitative two-dimensional gel electrophoresis of proteins. Methods Enzymol,1983,100:411～423

35 Bo F, Johan L, Gert A. Two-dimensional polyacrylamide gel electrophoresis of human lung cancer:qualitative aspects of tissue preparation in relation to histopathology. Electrophoresis, 1991,12: 509～515

36 Bo F, Alaiya AA, Auer G. Sample preparation of human tumors prior to two-dimensional electrophoresis of proteins. Electrophoresis, 1995,16:1087～1089

37 Seung OL, Sung-Jin P *et al*. Proteome Analysis of Hepatocellular carcinoma. Biochem Biophy Res Commol/Lunicat, 2002, 291: 1031～1037

38 Cecilia S, Daniela G *et al*. Renal cell carinoma and normal kidney protein expression. Electrophoresis, 1997,18:599～604

39 Reymond MA, Sanchez JC, *et al*. Standardized characterization of gene expression in human colorectal epithelium by two-dimensional electrophoresis. Electrophoresis, 1997,18:2842～2828

40 Reymond MA, Jean-Charles S *et al*. Specific sample preparation in colorectal cancer. Electrophoresis, 1997,18:622～624

41 Yuriy G, Jason S *et al*. Rolling circle amplification, a new approach to increase sensitivity for immol/Lunohistochemistry and flow cytometry. Am J Pathol,2001,159:63～69

42 Lizardi PM, Huang XH *et al*. Mutation detection and single- molecule counting using isothermal rolling-circle amplification. Nat Genet, 1998, 19:225～232

43 Darzynkiewicz Z, Bedner E *et al*. Laser-scanning cytometry: a new instrumentation with many applications. Exp Cell Res, 1999,249:1～12

44 Tarnok A, Gerstner AO. Clinical applications of laser scanning cytometry. Cytometry, 2002,50:133～143

45 Vignali DA. Multiplexed particle-based flow cytometric assays. J Immol/Lunol Methods, 2000,243:243～255

46 Michael R, Emmol/Lert-Buck *et al*. Increased gelatinase A(MMOL/LP-2) and cathepsin B activity in invasive tumor regions of human colon cancer samples. Am J Pathol,1994,145:1285～1290

47 Zhuang ZP *et al*. A microdissection technique for Archival DNA analysis of specific cell populations in lesions＜1mmol/L in size. Am J Pathol, 1995,146:620～625

48 Whetsell L *et al*. Polymerase chain reaction microanalysis of tumors from stained histological slides. Oncogene, 1992, 7:2355～2361

49 Radford DM *et al*. Allelic loss on chromosome 17 in ductal carcinoma *in situ* of the breast. Cancer Res, 1993,53: 294～295

50 Noguchi S *et al*. Clonal analysis of predominantly intraductal carcinoma and precancerous lesions of the breast by means of polymerase chain reaction. Cancer Res, 1994,54:21849～21853

51 Meier-ruge W, Bielser W *et al*. The laser in the Lowry technique for microdissection of freeze-dried tissue slices. Histochem J, 1976,8:387～401

52 Michael R, Emmol/Lert-Buck, Robert FB *et al*. Laser capture microdissection, Science, 1996,274:998～1001

53 Bonner RF, Emmol/Lert-Buck M *et al*. Laser capture microdissection: molecular analysis of tissue. Science, 1997, 278:1481～1483

54 Rosamode E, Banks MJ, Dunn MA *et al*. The potential use of laser capture microdissection to selectively obtain distinct populations of cells for proteomic analysis-preliminary findings. Electrophoresis, 1999,20:689～700

55 Michael R, Emmol/Lert-Buck, Gillespie JW *et al*. An approach to proteomic analysis of human tumors. Mol Carcinog, 2000,27:158～165

56 Zhou G, Li HM, DeCamp D *et al*. 2D differential in-gel electrophoresis for the identification of Esophageal Scans cell cancer-specific protein markers. Mol Cell Proteomics, 2002, 1.2:117~123
57 Karin S, Georgia L. Identification of expressed genes by laser-medicated manipulation of single cells. Nat Biotechnol, 1998, 16:737~742
58 Ludger F, Thomas K *et al*. Immol/Lunostaining for cell picking and real-time mRNA quantiation. Am J Pathol, 2000,147:1459~1466
59 Juha K, Lukas B *et al*. Tissue microarrays for high-throughput molecular profiling of toumor specimens. Nat Med, 1998, 4:844~847
60 Cloud PP, Lu C *et al*. Reverse phase protein microarrays which capture disease progression show activation of pro-survival pathways at the cancer invasion front. Oncogene, 2001,20: 1981~1989
61 Juan MG, Wang H *et al*. Protein based microarrays: a tool for probing the proteome of cancer cells and tissues. Proteomics, 2001,1:1279~1287
62 Richard M, Caprioli TB, Farmer J. Molecular imaging of biological samples: localization of peptides and proteins using MALDI-TOF MS. Anal Chem, 1997,69:4751~4760
63 Pierre C, Markus S, Caprioli RM. Direct profiling of proteins in biological tissue sections by MALDI-TOF MS. Anal Chem, 1999,71:5263~5270
64 Markus S, Pierre C *et al*. Imaging mass spectrometry: a new technology for the analysis of protein expression in mammol/Lalian tissues. Nat Med, 2001,7:493~496
65 Patrick G, Eckhard N, Hans L *et al*. Extraction of proteins from plant tissues for two-dimensional electrophoresis analysis. Electrophoresis,2003,24:207~216
66 Kose J. 2-D Proteome analysis protocols. Totowa: Humana Press, 1999,67~85
67 Kamo M *et al*. Separation and characterization of Arabidopsis thaliana proteins by two-dimensional gel electrophoresis. Electrophoresis, 1995, 16: 423~430
68 Economou A. Sec, durgs and rock'n'roll: antibiotic targeting of bacterial protein translocation. Emerg Ther Targets, 2001,5:141~153
69 Tjalsma H, Bolhuis A, Jongbloed JD *et al*. Signal peptide-dependent protein transport in Bacillus subtilis: a genome-based survey of the secretome. Microbiol Mol Biol Rev, 2000, 64:515~547
70 Economou A. Following the leader: bacterial protein export through the Sec pathway. Trends Microbiol, 1999, 7: 315~320
71 Dirk B, Sevil A, Meike W *et al*. Proteome analysis of secreted proteins of the gastric pathogen *Helicobacter pylori*. Infect Immun, 2002, July,p:3396~3403
72 Sauter ER, Zhu W, Fan XJ *et al*. Proteomic analysis of nipple aspirate fluid to detect biologic markers of breast cancer. Br J Cancer, 2002,86:1440~1443
73 Choe LH,Dutt MJ, Relkin N, *et al*. Studies of potential cerebrospinal fluid molecular marker for Alzheimer's disease. Electrophoresis,2002,23:2247~2251
74 Joshua NA, Susan MV *et al*. Toward a human blood serum proteome. Mol Cell Proteom,2002,1:947~955
75 Yuan X, Russell T, Wood G *et al*. Analysis of the human lumbar cerebrospinal fluid proteome. Electrophoresis, 2002, 23:1185~ 1196

第十四章　双向凝胶电泳

第一节　介　　绍

由于蛋白质组学是从整体水平研究细胞、组织、器官甚至是有机体(微生物)内的蛋白质表达变化,那么首先需要一种强有力的方法来分离、显示和分析成千上万种蛋白质。虽然目前蛋白质组学技术发展很快,如毛细管等电聚焦[1]、多维色谱[2]、微流芯片[3]等技术均已建立,但双向凝胶电泳仍然是目前蛋白质组学研究的主要分离方法。双向凝胶电泳的高分辨率、高重复性和兼具微量制备的性能是其他分离方法所不能匹敌的。

一、双向凝胶电泳简介

双向电泳的思路最早由 Smithies 和 Poulik[4]提出,蛋白质第一向电泳是根据其自由溶液迁移率(free-solution mobility),在滤纸条上进行;第二向电泳方向与第一向垂直,在淀粉胶上进行。随着聚丙烯酰胺凝胶的发明[5],双向电泳支持介质逐渐转向聚丙烯酰胺凝胶[6],即双向凝胶电泳。1969 年,双向凝胶电泳在原理上有了新的发展,以第一向为蛋白质等电聚焦的双向凝胶电泳建立起来(即 IEF-PAGE)[7,8]。20 世纪 70 年代初,第二向电泳中使用了十二烷基硫酸钠(SDS)[9~11],使第二向电泳基本上根据蛋白质的相对分子质量来分离,从而奠定了现代双向凝胶电泳的基础(即 IEF-SDS PAGE)。1975 年,O′Farrell[12]、Klose[13]和 Scheele[14]分别对双向凝胶电泳做进一步的优化,建立了高分辨的双向凝胶电泳。

双向凝胶电泳的基本原理是:先将蛋白质根据其等电点在 pH 梯度胶内进行等电聚焦,然后按照它们的相对分子质量大小进行 SDS-PAGE 第二次电泳分离。根据第一向等电聚焦条件和方式的不同,可将双向凝胶电泳分为三种系统:ISO-DALT、NEPHGE 和 IPG-DALT[15]。在 ISO-DALT 系统中,等电聚焦在聚丙烯酰胺管胶(tube gel)中进行,载体两性电解质 (carrier ampholyte)在外加电场作用下形成 pH 梯度。pH 梯度在碱性区不稳定(阴极漂移)、重复性不易掌握和上样量低是其主要缺点,并且一般不能分离等电点大于 8.0 的碱性蛋白质。优点是电泳设备要求不高,电泳溶液容易配制,易于开展工作。各种电泳条件经优化后,可达到非常高的分辨率,也可获得好的重复性[16,17],目前惟一分辨到 10 000 个蛋白质斑点的图谱正是采用了这一系统[16]。图 14.1 是我们利用 ISO-DALT 分离鼻咽癌细胞系 CNE-2 细胞总蛋白质的双向图谱。NEPHGE 为非平衡 pH 梯度电泳(nonequilibrium pH gradient electrophoresis),是 IPG(immobilized pH gradient)胶发明前用于分离碱性蛋白质的一种方法[18],蛋白质在等电聚焦场中达到平衡前结束电泳,第一向的 pH 也是依靠载体两性电解质和电场来建立。IPG-DALT 的第一向电泳是采用 1982 年发明的 IPG 胶[19],其 pH 梯度的形成依赖于不同 pK 值的一种合成的化合物 immobiline。该化合物是丙烯酰胺的衍生物,为非两性的弱酸和弱碱,在凝胶聚合过程中,能与聚

丙烯酰胺共价结合。因此,IPG 胶的 pH 梯度是稳定的,不依赖于外加电场,基本上克服了 ISO-DALT 的主要缺点[20],无论在重复性和上样量上均优于 ISO-DALT。

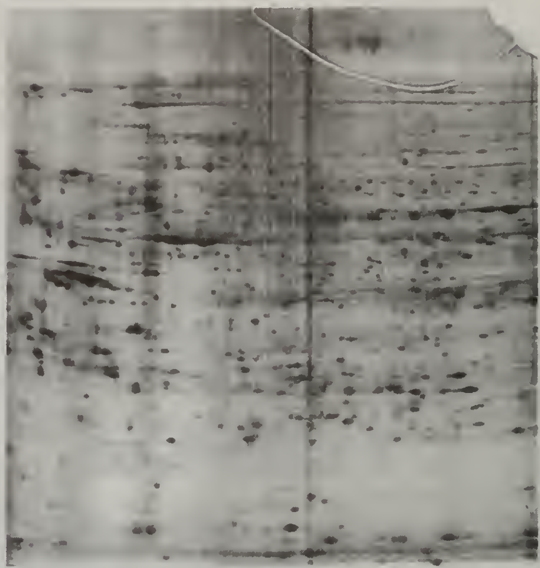


图 14.1 鼻咽癌细胞系 CNE-2 的双向凝胶电泳银染图谱

二、IPG-DALT 双向凝胶电泳

(一) 样品制备

要获得高分辨和高度重复的双向凝胶电泳图谱,蛋白质样品制备是极其关键的一步。首先,样品不能变质,要使用新鲜的样品或者将新鲜样品速冻(液氮或 - 70℃)起来也可以,样品需要处理时要尽量和处理前的保持一致。样品处理过程中注意防止污染,手套等必须戴上。不同类型的样品制备请见上一章。为了分离一种细胞或组织的蛋白质组,要尽可能使蛋白质组中的各种蛋白质成分溶解在裂解缓冲液中,整个蛋白质抽提过程要避免蛋白质的降解、修饰等。对于组织而言,由于组织的不均一性,保证组织细胞取样的重复性很重要,采用激光捕获微解剖(laser capture microdissection)技术可获得特定的一类细胞[29],但该技术目前在蛋白质组中的应用也存在局限性,主要的问题是每次所获得的细胞量太少。一般在样品裂解前和裂解中可结合机械破碎方法,如研磨、超声等。

裂解液的基本成分为尿素、去垢剂、还原剂和两性电解质等。尿素是一种优良的变性剂,通过断裂氢键和疏水键使蛋白质变性,增加其溶解性,而不影响蛋白质所带的电荷。尿素与硫脲结合使用使蛋白质的溶解性更好,尤其是疏水性的膜蛋白[30,31]。去垢剂有离子型、非离子型和兼性离子型等几类。离子型去垢剂,如 SDS 使蛋白质带上负电荷,干扰第一向等电聚焦而避免使用。但 Harder 等人在加裂解液抽提酵母细胞蛋白质之前,用 1%SDS 的缓冲液破碎细胞并加热可抑制相对分子质量较高的蛋白质被内源性蛋白水解酶降解,并提高了相对分子质量较高的蛋白质的溶解性[32]。传统的非离子型去垢剂,如 Triton X-100 和 NP-40,以前用得较多,现多数以兼性离子去垢剂 CHAPS(3-[(3-cholami-

dopropyl)-dimethylammonio]-1-propane sulfonate)代替。CHAPS 的终浓度可达到 4%。SB 3-10 (*N*-decyl-*N*, *N*-dimethyl-3-ammonio-1-propane sulfonate) 溶解疏水性氨基酸残基的能力更好,但它在高浓度的脲溶液中溶解性较差[33],终浓度应尽量小于 2%。最近,有一种新的兼性离子去垢剂为 ASB 14 (Amidosulfobetaine 14),Molloy 等[34]把有机溶剂抽提的大肠杆菌蛋白质溶于含 ASB 14 的增溶溶液中,发现蛋白质的溶解性比用 SB3-10 和 CHAPS 好;从 13 个被鉴定的蛋白质中发现 8 个是以前在双向胶上没有发现过的,其中有 5 个是新发现的疏水性蛋白质。还原剂多数使用 β-巯基乙醇、DTT 或 DTE(dithioerythreitol),通过打断二硫键使蛋白质彻底变性。Herbert 等人[35]用三丁基膦(tributyl phosphine, TBP)代替 DTT,在等电聚焦中大大增加了蛋白质的溶解性,并提高了蛋白质从第一向到第二向转移的效率。但 TBP 不溶于水,半衰期短,等电聚焦时可能并不起作用。Seow[36]等在肝癌样品处理时也发现 TBP 做还原剂时蛋白质的溶解性较差。两性电解质能促进蛋白质的溶解,吸附高浓度尿素在溶液中形成的氰酸盐离子(cyanate ions),离心时还有助于核酸的沉淀。此外,它还可以阻止样品蛋白质和 IPG 胶条中的固相化的两性电解质相互作用。样品中的核酸大分子会干扰等电聚焦和双向胶的银染,加入核酸水解酶或采用超声等方法可将其降解。为避免样品制备过程中蛋白质的降解,整个过程需在低温中进行(约 4℃),并可加入 PMSF 等蛋白酶抑制剂。Görg[36]总结了三种裂解液的配方:①9 mol/L Urea, 1% (*m*/*V*) DTT, 2%~4% (*m*/*V*) CHAPS, 2% (*V*/*V*) carrier ampholytes, pH3~10 和 10 mmol/L Pefabloc® Protease Inhibitor; ②7 mol/L Urea, 2 mol/L Thiourea,1% (*m*/*V*) DTT, 2%~4% (*m*/*V*) CHAPS, 2% (*V*/*V*) Carrier Ampholytes, pH3~10 和 10mmol/L Pefabloc® Protease Inhibitor; ③热的 SDS 缓冲液:1% (*m*/*V*) SDS,100mmol/L Tris-HCl, pH7.0,该缓冲液裂解样品后至少 3 倍稀释于①或②的裂解液中。强变性作用的裂解液中不加蛋白酶抑制剂亦可。

(二) 第一向:等电聚焦

1. IPG 胶条的重泡涨

商品化的胶条使用前是以干胶条的形式和塑料支撑薄膜粘在一起,加样前需要先撕去薄膜在重泡涨液中泡涨。重泡涨液的成分为 8 mol/L Urea, 2% CHAPS, 18mmol/L DTT, 0.5% IPG Buffer,痕量 Bromophenol Blue,和裂解液成分基本相同,在 10~100mmol/L 浓度范围内适当提高 DTT 的浓度可提高蛋白质的检测灵敏度[38]。重泡涨可以在等电聚焦盘内进行,此时样品已经加到重泡涨液中,在加样杯上样(cup loading)方式中,胶条在专门的重泡涨盘内泡涨完成后再转入等电聚焦盘内加样。对于非常碱性的窄 pH 范围的胶条,如 pH9~12 或 pH10~12,重泡涨液的成分要根据样品做适当修改。如核糖体蛋白,重泡涨液的成分为 8mol/L Urea, 10% (*V*/*V*) 2-propanol, 10% (*V*/*V*) Glycerol, 1% (*m*/*V*) CHAPS, 20 mmol/L DTT, 0.5% Pharmalyte 3—10, 0.2% (*m*/*V*) Methycellulose 或 16% (*V*/*V*) 2-propanol, 10% Glycerol, 1% (*m*/*V*) CHAPS, 20mmol/L DTT, 0.5% Pharmalyte。组蛋白的重泡涨液成分同上,但没有 CHAPS[36]。重泡涨时间至少要 10h,泡涨不充分会影响相对分子质量较高的蛋白质的吸收。重泡涨液的体积和胶条的长度相匹配,可参照产品说明书。重泡涨和等电聚焦过程中的温度都

稳定在20℃,温度太低尿素会结晶出来,温度不稳定也会造成胶上蛋白质点位置的变化[39]。重泡涨时加上30V或50V的低压会促进胶条对蛋白质的吸收,但这一点在制备型电泳时才有意义。

2. 加样

样品可以在重泡涨时加入,称为胶内重泡涨(in-gel rehydration),也可以重泡涨完成后再加,如加样杯上样。重泡后加样相对不可靠,操作比较困难而且蛋白质容易在胶和样品的界面产生沉淀。然而在某些特殊情况下,重泡涨后加样具有优势,如使用pH6～11或pH6～9的胶条分离碱性蛋白质时,在阳极用加样杯加样,可以得到更好的分离图谱。样品的胶内重泡涨是推荐的方式,可以增加蛋白质上样量,也避免了在阴极或阳极加样的方向问题。Sabounchi-Schutt[40]报道了一种新的加样方式,称其为纸桥加样(paper bridge loading),即胶条重泡涨好以后,用1.5cm×5.0 cm的滤纸做为纸桥覆盖在胶条的一端,将样品加到滤纸上,这样可显著提高蛋白质的上样量,并在分辨率和相对分子质量较高的蛋白质的分离方面有所改善,如图14.2所示,这种加样方式可以在Pharmacia公司24cm的等电聚焦盘上实现。分析型双向凝胶电泳上样量在几十到几百微克,制备型双向凝胶电泳上样量在数百(银染)到最大几十毫克(考马斯亮蓝染色)之间,上样量的大小和胶条的pH范围也有关系,pH范围越窄,上样量越大。由于蛋白质定量方法的重复性并不十分可靠,最佳的上样量要根据实验结果来优化。从我们的结果来看,分析型银染,Bio-Rad 17cm pH3～10L胶条的最佳上样量为60～80μg,Pharmacia 18cm pH3～10NL的最佳上样量为80～100μg;制备型银染,pH3～10胶条最大可加到400μg;制备型考马斯亮蓝染色,pH3～10胶条可加到1mg,pH4～7胶条可以加到1.5mg,在制备型电泳时为保证图谱质量减少变形,样品最好脱盐后再上样。

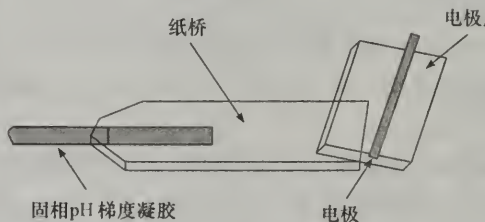


图14.2 纸桥加样法的示意图

3. 等电聚焦

重泡涨结束后,在胶条和电极的之间加上润湿的小纸片,它可以吸收盐离子和补充水分。在阴极加上20mmol/L DTT润湿的纸片可减少碱性端的水平条纹并且有助于碱性蛋白质的聚焦[41]。等电聚焦的电压上升分为三个阶段,首先加上一个比较低的电压,去除胶条中的过多的盐离子。然后开始加高压,电压上升的模式为缓慢上升。最后是持续高压,电压上升的模式为快速上升。Bio-Rad公司的第一向等电聚焦仪Protean IEF最多可加到10 000V的高压,Pharmacia公司IPGphor的最高电压为8 000V。Bio-Rad公司胶条总的电压时间积为7cm胶条8～35kVh,11cm胶条15～60kVh,17cm胶条25～100kVh。一般原则是,根据胶条的pH范围和上样量的多少,总电压时间积可以在一定范

表 14.1 Protean IEF 的聚焦条件设置

胶条长度 17cm;温度 20°C;最大限制电流 0.05mA per IPG strip;最大限制电压 10 000V	
分析型等电聚焦	
重泡涨：12h	
开始阶段的等电聚焦设置：	
250V 15min 电压快速上升	
10 000V 5h 电压快速上升	
稳定阶段的等电聚焦设置：	
pH3~10 胶条	pH4~7 胶条
10 000V~35 000Vh 电压快速上升	10 000V~60 000Vh 电压快速上升
制备型等电聚焦	
重泡涨：12h	
开始阶段的等电聚焦设置：	
250V 15min Rapid	
1 000V 1h Rapid	
4 000V 1h Rapid	
稳定阶段的等电聚焦设置：	
pH3~10 胶条	pH4~7 胶条
10 000V~80 000Vh 电压快速上升	10 000V~100 000Vh 电压快速上升



图 14.3 pH4~7 的人肝癌细胞系 MHCC97 的双向凝胶电泳银染图谱

围内调整。上样量大、pH 范围窄则总电压时间积高，分析型双向凝胶电泳限制电流可以放宽到 70 微安，制备型双向凝胶电泳的限制电流最好在 50 微安。等电聚焦完成后的胶条若不能及时进行第二向电泳，胶条可于 −20℃ 保存。以 17cm 胶条为例，优化的等电聚焦设置如表 14.1 所示，pH4～7 胶条的实验结果如图 14.3 所示。

（三）胶条的平衡

等电聚焦完成后，胶条必须经过平衡才能走第二向，平衡缓冲液的组分为 6mol/L 尿素，0.375mol/L Tris pH8.8，2% SDS，30% glycerol，2%（*m/V*）DTT 或 2.5%（*m/V*）IAA。平衡液的体积随胶条的长度而异，我们的用量为 17cm/18cm 胶条 3ml，11cm/13cm 胶条 2ml，大大低于说明书的用量。平衡液中尿素和甘油可增加溶液黏度，减少电内渗。电内渗作用是由固相化的两性电解质造成的，它影响蛋白质从第一向到第二向的转移。SDS 用于变性蛋白质，使蛋白质带上负电荷，这对第二向 SDS-PAGE 来讲是十分重要的。DTT 和 IAA 分别用于打开和封闭蛋白质的二硫键。溴酚蓝用来检测电泳过程。两次平衡的时间均为 15min，如果图谱的竖直条纹较多则平衡时间可适当延长到 20min，时间太短则蛋白质没有被 SDS 充分包裹，容易产生水平条纹。Galvani[42] 发现如果平衡液中有硫脲的话，它会迅速消耗 IAA 从而破坏其烷基化作用，并指出丙稀酰胺单体可以代替 IAA 且不会受到硫脲的干扰。我们用相同浓度的丙稀酰胺单体替代 IAA，验证了这一点。DTT 也可以用 5mmol/L TBP 代替。Lauber[43] 报道两次平衡可以简化为用 62.5 mmol/L Tris，2.3% SDS，5%巯基乙醇，pH6.8 的平衡液一次平衡 3.5min 即可，效果如何，我们没有试用。

（四）第二向 SDS-PAGE

第二向 SDS-PAGE 采用 Laemmli[44] 的不连续缓冲液系统，3T% 的 IPG 胶条可以看作为压缩胶，所以我们只使用分离胶，但也有人用压缩胶。IPG 胶条平衡好后，一般将胶条在电泳缓冲液里浸润一下，这样做一方面可以清洗胶条去除胶条上的平衡液，另一方面润湿的胶条也容易沿着玻璃板推到 SDS-PAGE 凝胶上端。IPG 胶条转移到 SDS-PAGE 胶有两种方式：一种方式是先将 IPG 胶条放到 SDS-PAGE 胶上，再加入熔化的琼脂糖溶液，另一种方式是先加入熔化的琼脂糖溶液，再将 IPG 胶条放到 SDS-PAGE 胶上。两种方式各有优劣。第一种方式图谱不易变形，但在两块胶之间容易有小气泡产生，第二种方式非常好地解决了胶之间的气泡问题，但琼脂糖容易凝固，插入 IPG 胶条时有时会造成图谱变形。相比之下，气泡对图谱的影响较小，故我们推荐第一种方式。SDS-PAGE 胶的浓度往往需根据研究目的和样品性质而定，最常用的是 12%和 12.5%的均一胶。采用均一胶的好处是灌胶方便，重复性好。但相对分子质量较高的蛋白质斑点较为聚集，分离不佳。采用 9%～16%的线性梯度胶，相对分子质量不同的蛋白质点在整块凝胶上分布比较均匀，如图 14.4 所示，但灌胶不易。电泳过程中温度一般控制在 10～15℃。



图 14.4　使用 9%～16%T 梯度胶的人肝癌细胞系
BEL-7404 的双向凝胶电泳银染图谱

第二节　试剂、设备和溶液配制

一、试　　剂

丙稀酰胺(Arc),甲叉双丙稀酰胺(Bis),三(羟甲基)氨基乙烷(Tris),甘氨酸,十二烷基磺酸钠(SDS),四甲基乙二胺(TEMED),过硫酸铵,二硫苏糖醇(DTT),碘代乙酰胺(IAA),尿素,3-[(3-胆酰胺丙基)-二乙铵]-丙磺酸(CHAPS),矿物油,均购自 Bio-Rad 公司,溴酚蓝为 Amersham Pharmacia 公司产品。预制 IPG 胶条及其 IPG 缓冲液为以上两家公司相应产品。所有缓冲液均用 Milli-Q 水配制。

二、仪　　器

超声细胞破碎仪(宁波新芝科仪研究所 JY92-2D);本实验室有两家公司的双向凝胶电泳实验仪器:Amesham-Pharmacia 公司的 IPGphor 水平电泳系统和 Hoefer SE 垂直电泳系统,如图 14.5 所示。

Protean IEF@ System 水平电泳系统和 Protean Ⅱ XL Cell 垂直电泳系统为 Bio-Rad 公司产品,如图 14.6 所示。



图 14.5 Amesham-Pharmacia 公司的 IPGphor 水平电泳系统
和 Hoefer SE 垂直电泳系统

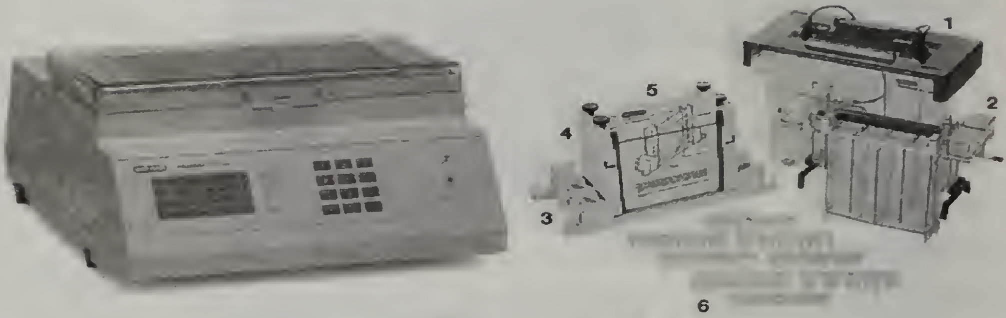


图 14.6 Bio-Rad 公司的 Protean IEF@ System 水平电泳系统和 PROTEAN II XL Cell 垂直电泳系统

三、溶液配制

1. 裂解液

8mol/L 尿素(Urea), 4% 3-[(3-胆酰胺丙基)-二乙铵]-丙磺酸(CHAPS), 40mmol/L 三(羟甲基)氨基乙烷(Tris), 65mmol/L 二硫苏糖醇(DTT) 40ml。

尿素	19.2 g
CHAPS	1.60g
Tris 碱	0.194g
DTT	0.401g

分装 -75℃保存。

2. 重泡涨液

8 mol/L 尿素, 2% CHAPS, 18mmol/L DTT, 0.5% IPG 缓冲液, 痕量溴酚蓝(Bromophenol blue trace) 40ml。

尿素	19.2g
CHAPS	0.8g
痕量溴酚蓝	

分装 -20℃保存。临用前加入 DTT 和 IPG buffer, 每 250μl 加 0.7mgDTT 和 1.25 μl IPG 缓冲液。

3. 平衡缓冲液

50mmol/L Tris-HCl pH8.8，6mol/L 尿素，30%甘油(Glycerol)，2%十二烷基磺酸钠(SDS)，痕量溴酚蓝

Tris-HCl pH8.8(1.5mol/L)	20ml
尿素	72.07g
甘油(87%，V/V)	69ml
溴酚蓝	痕量
水	to 200ml

临用前加入 DTT 和 IAA,浓度分别为 1%和 2.5%。

4. 0.5%的琼脂糖

0.1g 琼脂糖溶解在 20ml 1×电泳缓冲液,加入痕量溴酚蓝后,加热融化。于室温存放。

5. SDS-PAGE 贮存液和电泳缓冲液

(1) 丙烯酰胺(Acr)/甲叉双丙烯酰胺(Bis)(30%T, 2.67%C)

丙烯酰胺(Acrylamide)(Bio-Rad 161-0107)	146.0g
甲叉双丙烯酰胺(N,N-Methylene-bisAcrylamide)(Bio-rad161-0201)	4.0g

用 500ml 水溶解后,用 0.45μm 的微孔滤膜过滤,4℃避光保存,保存期为 30d。Arc、Bis 为神经性毒剂,配时请小心。

(2) 1.5mol/L Tris-HCl,pH8.8

54.45 g Tris 碱溶于 150 ml 水中,用盐酸调 pH 至 8.8,然后加水定容至 300ml。4℃保存。

(3) 10% (m/V)十二烷基磺酸钠(SDS)

10g SDS 溶于 60ml 水中,轻微振荡,然后加水定容至 100ml。SDS 很轻,称量时要小心,以免吸入。

(4) 10%过硫酸铵(Ammonium Persulfate)(m/V)

新鲜配制。

(5) 5×电泳缓冲液(1× =25mmol/L Tris, 192mmol/L 甘氨酸,0.1%SDS,pH8.3)

Tris 碱	45.0g
甘氨酸	216.0g
SDS	15.0g

溶于 3L 水中,pH8.3 不用调整。每次用时取 800ml,加水稀释到 4L。

第三节 实验操作

一、样品制备

以真核细胞为例,按 2×10^6细胞数加 150μl 体积的比例加裂解液。原核样品建议超

声裂解。真核生物样品可不经超声或温和超声。超声前预冷至4℃，超声时冰浴保持低温，时间为2min。注意勿产生泡沫。于4℃ 25 000*g*，离心1h，取上清。上清分装于-75℃保存。

二、蛋白质定量

在电泳前需要对样品的浓度进行测定，一般采用Bradford法。其他方法如BCA法、Biruet法和Lowry法定量蛋白质都是建立在将Cu^{+}还原为Cu^{2+}的颜色变化基础上的，而双向凝胶电泳裂解液中含有巯基还原剂，会影响这些定量方法的灵敏度。Bradford法是根据考马斯亮蓝G-250和蛋白质的结合来定量蛋白质的，但这种结合也会受到裂解液中的碱性Ampholytes和尿素的影响。

Ramagli[37]改进了Bradford法，先用0.1mol/L的盐酸中和样品，方法如下。

1. 染色液

Bio-Rad Protein Assay Dye Reagent Concentrate 500-0006，稀释5倍使用。或者自行配置：10 mg的考马斯亮蓝G250溶解于5ml的95%乙醇中，与10ml 85%磷酸混合定容到100ml。4℃保存，使用前用Whatman No.1滤纸过滤。

2. 标准蛋白质的配置

标准蛋白质使用卵白蛋白(ovalbumin)或γ球蛋白(γ-globulin)较好，牛血清白蛋白(BSA)做标准蛋白质时测得的样品浓度偏高，但在定量血清样品时比较准确。标准蛋白质以1mg/ml浓度溶于裂解液中，分装于-20℃保存。

3. 标准曲线的绘制

如表14.2所示。

表14.2　蛋白质定量的标准曲线绘制

标准蛋白质/μg	5	10	15	20	25	30	40	50
体积/μl	5	10	15	20	25	30	35	40
裂解液/μl	45	40	35	30	25	20	15	10
0.1mol/L HCl/μl	10	10	10	10	10	10	10	10
H_2O/μl	40	40	40	40	40	40	40	40

每管加3.5 ml染色液轻微摇匀，5 min后读取OD_{595}。

利用origin 5.0软件绘制的标准曲线如图14.7所示。

4. 样品浓度的测定

按比例稀释样品后按标准曲线的方法测定样品浓度。

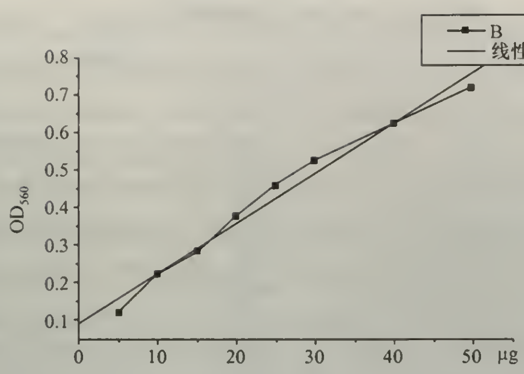


图 14.7　以牛血清白蛋白为标准蛋白质的蛋白质定量曲线

三、双向凝胶电泳

1. 加样

取出重泡涨液解冻，以 13cm 胶条为例，将 30～100μg 样品和重泡涨液充分混合，总体积为 250L。低速离心后，加入持胶槽，取出胶条撕去表面膜，胶面朝下从一端小心放入持胶槽，使重泡涨液浸润整个胶条，注意不要有气泡，胶条尖端和持胶槽尖端匹配，用矿物油覆盖。（考马斯亮蓝显色样品量为 100～300μg，制备型样品量为 1～10mg。）

2. 等电聚焦

重泡涨 12h；500V 1h；1 000V 1h；8 000V 6h。整个过程在 20℃ 进行，电流上限为 70mA。结束后记下总电压小时和电流。胶条可在 －20℃ 中保存或立即平衡。

3. 平衡

胶条取出后，先用滤纸吸去附着的矿物油，然后在 1％DTT 的平衡缓冲液中，置于摇床上平衡 15min 后，然后转入 2.5％IAA 平衡缓冲液中平衡 15min。2ml 平衡液完全够用，平衡时胶条可以放在 Bio-Rad 或 Amesham-Pharmacia 公司的重泡涨盘里进行。

4. SDS-PAGE

(1) 灌胶

按表 14.3 配制，抽气 15min 后加入 SDS、10％AP、TEMED。

凝胶时为保证每次灌出的胶高度一致，最好用尺量一下，一般距离玻璃板顶端 0.5cm，用水或水饱和的异丁醇封胶。凝胶聚合后用电泳缓冲液或水冲洗一下胶上端，除去未聚合的单体，同时也使胶条容易推到 SDS-PAGE 胶上端。

(2) 第一向胶条的转移

用镊子夹住一向胶条的一端，放到玻璃板上，胶面朝外，然后用平头镊子将小心地将胶条推到二向胶的上端。两块胶之间，经常会出现细小的气泡，可以用镊子调整一下胶条，尽量除去气泡，然后加入融化的 0.5％的琼脂糖封胶，此时尽量少动一向胶条。

表 14.3 SDS-PAGE 的凝胶配制

聚丙烯酰胺单体(%T,2.67%C)	分离胶	
	12.5%T	X%T
水	31.78ml	V-all the following
Arc/Bis (30%T, 2.67%C)	41.67ml	X * V/30ml
1.5mol/L Tris-HCl, pH8.8	25ml	V/4ml
10%SDS	1.0ml	V/100ml
10% 过硫酸铵	500μl	V/200ml
TEMED	50μl	V/2000ml
总体积	100ml	

(3) 电泳条件设置如表 14.4 所示

表 14.4 第二向 SDS-PAGE 电泳的设置

1.0mm Spacer	15mA/gel 走 15min ,然后 30mA/gel 走至溴酚蓝离胶下沿 0.5cm
1.5mm Spacer	20mA/gel 走 15min,然后 40mA/gel 走至溴酚蓝离胶下沿 0.5cm

5. 干胶

(1) 手工干胶

多孔聚乙烯膜两张,剪成比胶稍大的尺寸,浸泡在水中。取出一张在玻璃板(如灌胶用的玻璃板)上铺平,将凝胶平铺其上,胶上再覆盖一层膜,做成“三明治”夹饼状。四周用有机玻璃条压住,再用夹子夹紧。注意凝胶与两层膜之间不得有气泡,膜表面尽量保持光滑平整。于室温干燥,一般 24～72h 内即可干燥,时间差不多时注意观察,否则胶很容易干裂。



图 14.8 Bio-Rad 公司的干胶机

(2) 干胶机干胶

先将干胶机的滤纸板用水润湿,再将膜和胶放到滤纸板上,操作和注意事项基本同上。盖上橡胶垫,打开抽气装置后,检查橡胶垫四周是否已经吸牢,然后再盖上干胶机盖子,一般 80℃ 1h 即可。Bio-Rad 的干胶机如图 14.8 所示,抽气一般用水泵,该公司也有配套的抽气装置。

第四节 胶上蛋白质的检测

一、介　　绍

显示双向凝胶上的蛋白质点的方法有多种,例如,考马斯亮蓝 R-250 (CBB R-250)、CBB G-250、酰胺黑(Amido Black)、丽春红 S(Ponceau S)、银染、负染、荧光染料染色以及

放射性同位素标记等。放射性同位素标记法显示蛋白质的灵敏度非常高，同位素的掺入是以细胞代谢为基础的，不同蛋白质掺入同位素如^{35}S的量不同，因此，有时难以反映蛋白质组中各种蛋白质的实际丰度，从胶上切取蛋白质斑点的不便也是影响它被广泛使用的一个因素，但用来监测蛋白质组表达变化时有其潜力。银染和考马斯亮蓝染色是蛋白质组研究中最为广泛使用的两种染色方法，银染有 100 多种不同的方法，分为两大类，碱性银氨染色法和酸性硝酸银染色法。酸性硝酸银染色法和碱性银氨染色法相比，它对酸性蛋白质的染色灵敏度稍高，但对碱性蛋白质的染色灵敏度稍低。酸性硝酸银染色法的原理是一个照相的过程，凝胶在酸性环境下与硝酸银温育，银离子附于蛋白质表面，然后在碱性环境中里用福尔马林还原成金属银。碱性银氨染色法的原理是用氨水来形成银氨复合物，银氨复合物与胶内 SDS-蛋白质复合物结合，然后在酸性条件下用福尔马林将银氨还原成金属银。由于氨的不稳定性，酸性硝酸银染色法更适合于定量比较分析。一般认为银染的灵敏度是能染 1～10ng 的蛋白质，CBB R-250 的灵敏度在 100ng 左右，CBB G-250 的灵敏度稍高于 CBB R-250[45]。酰胺黑较广泛地用于染色电转移到膜上的蛋白质，其检测限在 30～50ng[45]。负染中的锌染[46]和铜染[47]是近年来逐渐受到重视的染色方法，染色后凝胶背景呈不透明的乳白色或青绿色，而蛋白质斑点或条带则是透明的。我们比较了银染、CBB R-250、锌染和铜染对蛋白质的染色灵敏度，银染能检测到 8ng 的肌红蛋白；CBB R-250 能检测到 62ng 的肌红蛋白和同量的牛血清白蛋白；铜染对肌红蛋白的灵敏度稍高于 CBB R-250，为 31ng，但只能检测到 125ng 的牛血清白蛋白条带；锌染的灵敏度与银染相当，8ng 的肌红蛋白或牛血清白蛋白均能被检测出来。CBB R-250、CBB G-250 和锌染对蛋白质的后续分析有较好的兼容性。由于金属负染蛋白质是可逆的，具有高灵敏度的锌染在蛋白质组研究中会具较大潜力，但对手工切胶而言，从锌染的凝胶上切取蛋白质斑点要困难些，采用自动切胶机可解决这一问题，图 14.9 是用光密度仪（densitometer）扫描锌染的双向凝胶得到的图谱。银染由于其较高的灵敏度，多数用来作为图谱分析为目的的染色方法，但通过染色方法的改进，银染后的蛋白质也可用来分析鉴定[48]。Scheler 等人[49]用不同的染色方法，结合酶解和质谱分析双向凝胶电泳分离的蛋白质，结果表明 CBB G-250 和锌染对量低的蛋白质斑点能得到较好的识别鉴定；与之相比，CBB



图 14.9　人肝癌细胞系 MHCC97 的部分双向凝胶电泳锌染图谱

R-250 染色量大的蛋白质斑点，酶解和质谱分析后可达 90％的序列覆盖率，对量中等的蛋白质斑点也可达 80％的覆盖率。染色方法的确定一般要根据研究需要和后续分析的灵敏度而定。

利用荧光试剂显影凝胶分离蛋白质在蛋白质组研究中的应用越来越广泛。传统的荧光标记往往需要共价修饰蛋白质的自由氨基、羧基或巯基[50, 51]，这种共价修饰的荧光标记法有许多缺点：等电聚焦前标记蛋白质会改变蛋白质的等电点，从而引起蛋白质的不正常迁移[50, 51]；不同蛋白质中荧光基团可修饰的功能基团不同，因此不同蛋白质对标记的灵敏度不同；荧光基团的共价修饰会干扰蛋白质的分离后分析，如蛋白质相对分子质量的轻微改变影响质谱识别鉴定蛋白质。然而，非共价修饰的荧光染料如 sypro red 和 sypro orange 等则与前者不同，它有许多优点。sypro red 和 sypro orange 对聚丙烯酰胺凝胶中蛋白质染色非常灵敏[45]，能检测 1～10ng 的蛋白质，可与银染媲美，且染色的动态范围比银染大，达 3 个数量级；染色步骤简单，一步染色，无需脱色；染色后的蛋白质适于 Edman 降解测序和质谱分析。但图像采集时需 UV 光源或激光扫描仪。一种新型的荧光染色试剂 sypro ruby，其染色原理是利用重金属铷的螯合物来和胶上蛋白质结合，染色专一性好，灵敏度高，而且不会担心染色过度。Lopez[52]比较了 sypro ruby 荧光染色和银染两种方法，前者的定量重复性(相关系数＝0.9)要比后者(相关系数＝0.8)稍好，银染的线性动态范围为 8～60ng，而 sypro ruby 染色的线性动态范围为 1～2ng 到 1～2μg，sypro ruby 的线性动态范围要比银染高 700 倍左右，而且 sypro ruby 染色在蛋白质酶解的序列覆盖率和质谱图信噪比方面都优于银染。图 14.10 是我们用 sypro ruby 染色得到的效果图。

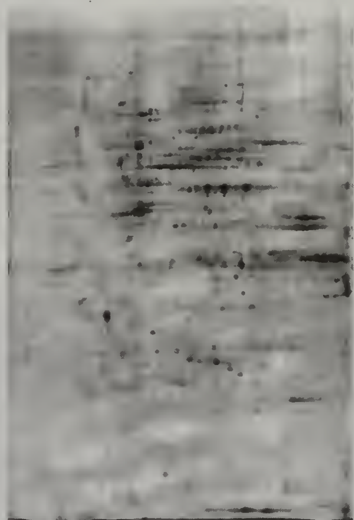


图 14.10　人肝癌细胞系 MHCC97 的 sypro ruby 染色图谱(Bio-Rad 11cm pH3～10 胶条)

观察功能相关的蛋白质表达的平行性变化是差异蛋白质组(differential proteome)研究的主要内容，实验需要多次重复，大量双向凝胶图像需要用计算机软件一张一张地分析、比较，使用改进的染色方法或图谱成像方式不仅能简化图像分析，并且可以减少实验工作量。Ünlü 等[53]建立了一种称为差异凝胶电泳(differential gel electrophoresis, DIGE)

的方法，实际上是采用两种荧光试剂分别标记两个样品，然后混在一起进行双向凝胶电泳。他们用两种花青染料(cyanine dyes)丙基-Cy3和甲基-Cy5分别标记经IPTG(isopropylthiogalacto-pyranoside)和未经IPTG诱导的大肠杆菌蛋白质抽提物，两样品混合后在一块胶上双向凝胶电泳，因两种染料分别具有不同的激发光和发射光波长，从一块凝胶上可以采集到代表两种样品的两张图像，经计算机处理，得到一张Cy3/Cy5的比例图(ratio image)(相当于差减图)，可以很清楚地看到GAL4VP16蛋白被IPTG诱导。图14.11是模式系统里用Cy3/Cy5染色得到的结果。

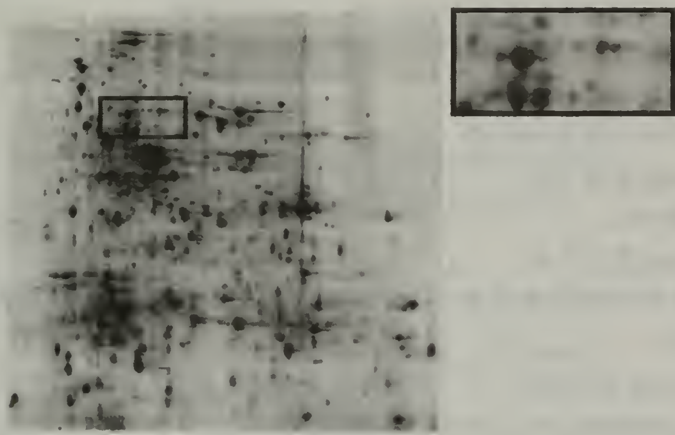


图14.11 以Cy3和Cy5为染料的模式系统的双向-差异凝胶电泳图谱

蓝色为Cy3的颜色，红色为Cy5的颜色。(AmershamPharmacia公司提供)

由Bernhardt等[54]建立的双通道成像(dual channel imaging)是另一种在一张电泳图谱上显示蛋白质组表达变化的方法。在细菌用L-[^{35}S]甲硫氨酸脉冲标记5 min后裂解，提取蛋白质，双向凝胶电泳后凝胶经过银染，采集银染的双向图谱，同一块胶再用PhosphorImager采集同位素标记蛋白质的双向图谱。银染图谱经软件转换为绿色的图谱(即保留绿色通道)，同位素图谱转换为红色图谱，然后将两图谱合成一张图谱，可以看到新合成的蛋白质为红色斑点，细胞内早已合成的蛋白质为绿色斑点，一直在合成的为黄色斑点。这种方法免去了凝胶配比的麻烦，但操作过程中应戴干净的手套以避免同位素污染。

二、染色方法

(一) 考马斯亮蓝染色

(1) 常规考马斯亮蓝染色

①滤纸过滤染色液[50%甲醇，10%乙酸，0.25% (m/V) CBB R350]。

②水洗胶5min。

③染色液中染色3h。

④5%(V/V)乙酸，10%(V/V)甲醇溶液脱色，反复更换脱色液直至蛋白质点在背景上变得清晰。

(2)胶体考马斯亮蓝染色[55]

①12% TCA 固定 1h。

②0.1%（m/V）CBB G250,2%（m/V）磷酸,10%过硫酸铵,20%甲醇染色 10h,注意染色盒要密封。

③25%甲醇溶液洗 1min,重复 2 遍。胶可置于 20%过硫酸铵中保存。

（二）银染

(1) 分析型银染(不能用于质谱鉴定)[56]

①水洗胶 5min;

②40%乙醇,10%乙酸固定 1h。

③5%乙醇,5%乙酸固定过夜。

④水洗 20min,重复 2 遍。

⑤5%戊二醛 1h。

⑥水洗 30min,重复 4 遍。

⑦250ml 新鲜配制的银氨液(1ml NaOH 与 5ml 饱和氨水混合再加 10ml 20% $AgNO_3$)/胶 45min。

⑧水洗 3min,重复 3 遍。

⑨1.5ml 1%柠檬酸、125μl 37%甲醛/胶显色。

⑩5%乙酸停显 10min。

⑪水洗 10min,重复 3 遍。

(2) 快速银染(和质谱鉴定相匹配)[57]

①水洗 5min。

②40%乙醇、10%乙酸重复固定 15min,重复 2 遍。

③30%乙醇、0.2%硫代硫酸钠、6.8%乙酸钠 30min。

④水洗 5min,重复 3 遍。

⑤2.5%碳酸钠,0.04%的福尔马林(36%的甲醛溶液)显影至背景出现。

⑥1.46%EDTA 溶液停显 10min。

⑦水洗 10min,重复 3 遍。

（三）锌染[46]

①水洗 1min。

②0.2mol/L 咪唑、0.1%SDS 溶液 10min。

③0.2mol/L 氯化锌 1min 至胶背景变为纯白。

④水洗 5s,重复 3 遍。

（四）ruby 荧光染色[43]

①10%甲醇、7%乙酸固定 1h。

②700ml sypro ruby 溶液/胶染色 3h。

③10%甲醇、7%乙酸洗 1h。

以下简述双向凝胶电泳的问题与原因[58~60]，如表 14.5 至表 14.7 所述。

表 14.5　第一向等电聚焦的问题

问题与现象	可能原因和解决措施
IPG 胶条在靠近电极附烧焦	电流限制太高，最大电流限制设为 50μA IPG 胶条没有完全重泡涨；应检查重泡涨液的体积，保证重泡涨液的体积足够 IPG 胶条过于干燥；重泡涨和等电聚焦时加矿物油
电压太低，达不到最高电压	样品盐浓度太高 脱盐或稀释样品，等电聚焦 1～2h 后更换电极垫片。盐浓度不得超过 100mmol/L，最好控制在 10mmol/L 以内 等电聚焦的设置错误 检查电压设置
电流为零或很低	持胶槽和电极接触不好，或者电极和胶条的接触不好 检查以上接触情况，检查胶条是否完全重泡涨
尿素在胶条表面结晶	温度太低 温度控制设为 20℃，室温不能太低

表 14.6　第二向 SDS-PAGE(垂直系统)的问题

问题与现象	可能原因和解决措施
开始时没有电流	上下槽是否有电泳缓冲液
电泳太慢	检查电泳缓冲液是否配错；检查分离胶缓冲液是否配错
溴酚蓝前沿一边上翘	阴极电泳缓冲液没有接触胶的一端；保证电泳缓冲液覆盖了胶的全长
溴酚蓝前沿两边上翘	冷凝不好；检查冷凝器是否工作正常
溴酚蓝前沿不规则	受 IPG Buffer 影响，对最终结果没有影响 胶条太旧或二向成品胶太旧；检查日期是否过期 IPG 胶条转移到二向时下面有气泡；保证胶条和二向胶顶部紧密接触，无气泡

表 14.7　染色的问题

水平条纹	聚焦时间太短(尤其对于相对分子质量较高的蛋白质)或太长(蛋白质会分解)； 优化聚焦时间
	去垢剂浓度太低或使用了不合适的去垢剂； 检查去垢剂浓度，测试不同的去垢剂
	尿素浓度太低； 重泡涨液的尿素浓度不低于 8mol/L
	IPG 胶条时间太短，没有重泡涨到其原始厚度； 重泡涨胶条至 0.5mm 厚，时间大于 6h 或过夜
	裂解液或重泡涨液的 DTT 浓度不够； 裂解液的 DTT 浓度为 65mmol/L 或重泡涨液的 DTT 浓度为 18mmol/L
	单个蛋白质的多重氧化； 加足够量 DTT，等电聚焦时加上矿物油，矿物油脱气或充上惰性气体

续表

水平条纹	由于 DTT 朝阳极移动 碱性端 DTT 消耗； 在阴极加上 20mmol/L DTT 润湿的电极垫片
	错误的样品加样区域； 检查阳极还是阴极加样比较好，或者使用样品的胶内重泡涨方式加样
	由于样品的内源性水解造成的假象； 用 TCA 丙酮法处理失活蛋白酶，加入蛋白酶抑制剂或与 SDS 一起煮
	空气中的 CO_2 干扰； 等电聚焦时加矿物油，密封电泳腔或在腔内加上 NaOH 润湿的纸片去除 CO_2
	相对分子质量为 68 000 或 55 000 的条纹，可能有角蛋白或白蛋白污染，或者由巯基乙醇造成； 过滤所有溶液，用 DTT 取代巯基乙醇
	蛋白质加样处产生沉淀； 等电聚焦时采用低电压
	蛋白质抽提时有不溶物在样品中； 高速离心 40 000*g* 1h
垂直条纹	蛋白质从一向转移到二向时发生沉淀； 提高平衡的 SDS 浓度，适当延长平衡时间
	蛋白质上样量过大； 降低上样量
	水不够纯净，SDS-PAGE 胶配制时没有过滤（背景有极细的条纹）或者胶条平衡后仍带有额外的 DTT； 使用 Millipore 水，过滤胶液。使用 IAA 平衡 IPG 胶条，去除额外的 DTT
	成对的竖直条纹遍布于胶上； 可能是由硫脲引起的，原因尚不明

（丁士健）

参 考 文 献

1 Tang W, Harrata AK, Lee CS. Two-dimensional analysis of recombinant *E. coli* proteins using capillary isoelectric focusing electrospray ionization mass spectrometry. Anal Chem, 1997,69:3177～3182

2 Link AJ, Eng J, Schieltz DM *et al*. Direct analysis of protein complexes using mass spectrometry. Nat Biotech, 1999, 17:676～682

3 Pinto DM, Ning Y, Figeys D. An enhanced microfluidic chip coupled to an electrospray Qstar mass spectrometer for protein identification. Electrophoresis, 2000,21:181～190

4 Smithies O, Poulik MD. Two-dimensional electrophoresis of serum proteins. Nature, 1956,177:1033

5 Raymond S, Weintraub L. Acrylamide gel as a supporting medium for zone electrophoresis. Science, 1959,30:711

6 Margolis J, Kenrick KG. Two-dimensional resolution of plasma proteins by combination of polyacrylamide disc and gradient gel electrophoresis. Nature, 1969,221: 1056～1057

7 Dale G, Latner AL. Isoelectric focusing of serum proteins in acrylamide gels followed by electrophoresis. Clin Chim Acta, 1969,24:61～68

8 Macko V, Stegemann H. Mapping of potato proteins by combined electrofocusing and electrophoresis identification of varieties. Hoppe-seyler's Z Physiol Chem, 1969,350:917～919

9 Stegemann H. Proteinfraktionierungen in polyacrylamid und die anwendung auf die genetische analyse bei pflanzen. Angew Chem, 1970,82:640

10 Martini OHW, Gould HJ. Enumeration of rabbit reticulocyte ribosomal proteins. J Mol Biol,1971,62:403～405

11 Barrett T, Gould HJ. Tissue and species specificity of non-histone chromatin proteins. Biochim Biophys Acta, 1973, 294:165～170

12 O'Farrell PH. High resolution two-dimensional electrophoresis of proteins. J Biol Chem, 1975,250:4007～4021

13 Klose J. Protein mapping by combined isoelectric focusing and electrophoresis of mouse tissues. A novel approach to testing for induced point mutations in mammals. Humangenetik,1975,26:231～243

14 Scheele GA. Two-dimensional gel analysis of soluble proteins. J Biol Chem, 1975,250:5375～5385

15 Humphery-Smith I, Cordwell SJ, Blackstock WP. Proteome research: complementarity and limitations with respect to the RNA and DNA words. Electrophoresis, 1997,18:1217～1242

16 Klose J, Kobalz U. Two-dimensional electrophoresis of proteins: an undated protocol and implications for a functional analysis of the genome. Electrophoresis, 1995,16:1034～1059

17 Lopez MF, Patton WF. Reproducibility of polypeptide spot positions in two-dimensional gels run using carrier ampholytes in the isoelectric focusing dimension. Electrophoresis, 1997,18:338～343

18 O'Farrell PZ, Goodman HM, O'Farrell PH. High resolution two-dimensional electrophoresis of basic as well as acidic proteins. Cell, 1977,12:1133～1142

19 Bjellqvist B, Ek K, Righetti PG *et al*. Isoelectric focusing in immobilized pH gradients: principle, methodology and some applications. J Biochem Biophys Methods, 1982,6:317～339

20 Görg A, Postel W, Günther S. The current state of two-dimensional electrophoresis with immobilized pH gradients. Electrophoresis, 1988,9:531～546

21 Bjellqvist B, Pasquali C, Ravier F *et al*. A nonlinear wide-range immobilized pH gradient for two-dimensional electrophoresis and its-definition in a relevant pH scale. Electrophoresis, 1993,14:1357～1365

22 Bjellqvist B, Sanchez JC, Pasquali C *et al*. Micropreparative two-dimensional electrophoresis allowing the separation of samples containing milligram amounts of proteins. Electrophoresis, 1993,14:1375～1378

23 Rabilloud T, Valette C, Lawrence JJ. Sample application by in-gel rehydration improves the resolution of two-dimensional electrophoresis with immobilized pH gradients in the first dimension. Electrophoresis, 1994,15:1552～1558

24 Langen H, Roder D, Juranville JF *et al*. Effect of protein application mode and acrylamide concentration on the resolution of protein spots separated by two-dimensional electrophoresis. Electrophoresis, 1997,18:2085～2090

25 Gorg A, Obermaier C, Boguth G *et al*. Recent developments in two-dimensional gel electrophoresis with immobilized pH gradients: wide pH gradients up to pH 12, longer separation distances and simplified procedures. Electrophoresis, 1999,20:712～717

26 Wasinger VC, Bjellqvist B, Humphery-Smith I. Proteomic 'contigs' of Ochrobactrum anthropi, application of extensive pH gradients. Electrophoresis, 1997,18:1373～1383

27 Urquhart BL, Atsalos TE, Roach D *et al*. 'Protemic contigs' of Mycobacterium tuberculosis and Mycobacterium bovis (BCG) using novel immobilised pH gradients. Electrophoresis, 1997,18:1384～1392

28 Adessi C, Miege C, Albrieux C *et al*. Two-dimensional electrophoresis of membrane proteins: a current challenge for immobilized pH gradients. Electrophoresis, 1997,18:127～135

29 Banks RE, Dunn MJ, Forbes MA *et al*. The potential use of laser capture microdissection to selectively obtain distinct populations of cells for proteomic analysis-preliminary findings. Electrophoresis, 1999,20:689～700

30 Pasquali C, Fialka I, Huber LA. Preparative two-dimensional gel electrophoresis of membrane proteins. Electrophoresis, 1997,18:2573～2581

31 Rabilloud T, Adessi C, Giraudel A *et al*. Improvement of the solubilization of proteins in two-dimensional electrophoresis with immobilized pH gradients. Electrophoresis, 1997,18:307～316

32 Harder A, Wildgruber R, Nawrocki A *et al*. Comparison of yeast cell protein solubilization procedures for two-dimensional electrophoresis. Electrophoresis, 1999,20:826～829

33 Herbert B. Advances in protein solubilisation for two-dimensional electrophoresis. Electrophoresis, 1999,20:660～663

34 Corthals GL, Molloy MP, Herbert BR *et al*. Extraction of Escherichia coli proteins with organic solvents prior to two-dimensional electrophoresis. Electrophoresis, 1999,20:701～704

35 Herbert BR, Molloy MP, Gooley AA *et al*. Improved protein solubility in two-dimensional electrophoresis using tributyl phosphine as reduing agent. Electrophoresis, 1998,19:945～951

36 Seow TK, Ong SE, Liang RC *et al*. Two-dimensional electrophoresis map of the human hepatocellular carcinoma cell line, HCC-M, and identification of the separated proteins by mass spectrometry. Electrophoresis, 2000,21:1787～1813

37 Ramagli LS. 2-D Proteome Analysis Protocols. In: Link AJ ed. Methods in Molecular Biology. Vol 112, Totowa: Humana Press, 1999, 99～103

38 2-D Electrophoresis for proteomics: a methods and product manual. Bio-Rad Company

39 Gorg A, Postel W, Friedrich C *et al*. Temperature-dependent spot positional variability in two-dimensional polypeptide patterns. Electrophoresis, 1991,12:653～658

40 Sabounchi-Schutt F, Astrom J, Olsson I *et al*. An immobiline DryStrip application method enabling high-capacity two-dimensional gel electrophoresis. Electrophoresis, 2000,21:3649～3656

41 Gorg A, Boguth G, Obermaier C *et al*. Two-dimensional polyacrylamide gel electrophoresis with immobilized pH gradients in the first dimension (IPG-Dalt): the state of the art and the controversy of vertical versus horizontal systems. Electrophoresis, 1995,16:1079～1086

42 Galvani M, Rovatti L, Hamdan M *et al*. Protein alkylation in the presence/absence of thiourea in proteome analysis: a matrix assisted laser desorption/ionization-time of flight-mass spectrometry investigation. Electrophoresis, 2001,22: 2066～2074

43 Lauber WM, Carroll JA, Dufield DR. Mass spectrometry compatibility of two-dimensional gel protein stains. Electrophoresis, 2001,22:906～918

44 Laemmli UK. Cleavage of structural proteins during the assembly of the head of bacteriophage T4, Nature, 1970,22: 680～685

45 Patton WF. Proteome analysis. Ⅱ. Protein subcellular redistribution: linking physiology to genomics via the proteome and separation technologies involved. J Chromatogr B, 1999,722: 203～223

46 Fernandez-Patron C, Calero M, Collazo PR *et al*. Protein reverse staining: high-efficiency microanalysis of unmodified proteins detected on electrophoresis gels. protein reverse staining: high-efficiency microanalysis of unmodified proteins detected on electrophoresis gels. Anal Biochem, 1995,224:203～211

47 Lee C, Levin A, Branton D. Copper staining: a five-minute protein stain for sodium dodecyl sulfate-polyacrylamide gels. Anal Biochem, 1987,166:308～312

48 Shevchenko A, Wilm M, Vorm O *et al*. Mass spectrometric sequencing of proteins from silver-stained polyacrylamide gels. Anal Chem, 1996,68:850～858

49 Scheler C, Lamer S, Pan Z *et al*. Peptide mass fingerprint sequence coverage from differently stained proteins on two-dimensional electrophoresis patterns by matrix assisted laser desorption/ionization-mass spectrometry (MALDI-MS). Electrophoresis, 1998,19:918～927

50 Urwin V, Jackson P. A multiple high-resolution mini two-dimensional polyacrylamide gel electrophoresis system: imaging two-dimensional gels using a cooled charge-coupled device after staining with silver or labeling with fluorophore. Anal Biochem, 1991,195:30～37

51 Urwin V, Jackson P. Two-dimensional polyacrylamide gel electrophoresis of proteins labeled with the fluorophore monobromobimane prior to first-dimensional isoelectric focusing: imaging of the fluorescent protein spot patterns using a cooled charge-coupled device. Anal Biochem, 1993,209:57～62

52 Lopez MF, Berggren K, Chernokalskaya *et al*. A comparison of silver stain and SYPRO Ruby protein stain with re-

spect to protein detection in two-dimensional gels and identification by peptide mass profiling. Electrophoresis, 2000, 21:3673～3683

53 Unlu M, Morgan M, Minden JS. Difference gel electrophoresis: a single gel method for detecting changes in protein extracts. Electrophoresis, 1997,18:2071～2077

54 Bernhardt J, Buttner K, Scharf C *et al*. Dual channel imaging of two-dimensional electropherograms in Bacillus subtilis. Electrophoresis, 1999,20:2225～2240

55 Neuhoff V, Arold N, Taube D *et al*. Improved staining of proteins in polyacrylamide gels including isoelectric focusing gels with clear background at nanogram sensitivity using Commassie Brilliant Blue G-250 and R-250, Electrophoresis, 1988,9:255～262

56 俞利荣，王楠，夏其昌等. 人肝癌细胞系 BEL-7404 和正常肝细胞系 L-02 表达的蛋白质组双向凝胶电泳分析. 科学通报，2000,45:170～178

57 Yan JX, Robin W, Tom B *et al*. A modified silver staining protocol for visualization of proteins compatible with matrix-assited laser desorption/ionization and electrospray ionization mass spectrometry. Electrophoresis, 2000,21:3666～3672

58 2-D Electrophoresis for proteomics: a methods and product manual. Bio-Rad Company

59 2-D Electrophoresis using immobilized pH gradients: principles and methods. Amersham Pharmacia Company

60 http://www.weihenstephan.de/blm/deg/

第十五章 电泳图谱的图像分析

第一节 介 绍

双向凝胶电泳根据等电点和相对分子质量的不同可一次分离数千种蛋白质，经染色后蛋白质在聚丙烯酰胺凝胶上形成密度不同、分布不均的复杂点图谱。如何对图谱上的这些蛋白质点进行检测、定量、比较、分析和归类是双向凝胶电泳分析软件所要解决的问题。

双向凝胶电泳图谱分析一般包括以下几个步骤：图像采集、图像加工、点的检测和定量、凝胶匹配、数据分析、数据解释和构建数据库。图像采集是通过扫描仪或数码相机将凝胶转换为数字图像，其参数设置对于以后的蛋白质点的内在信息抽提是比较重要的，因为以后的整个图像加工过程实质上是给图谱上的每个蛋白质点进行精确的定位和定量(以面积和密度作为指标)。图像加工主要是指软件通过一定的算法来去除图谱上非蛋白质点的杂质和染色不同造成的不同深浅的背景。点的检测主要是产生每个点的 X、Y 轴坐标，外形参数和密度参数，并对每个蛋白质点做数字模拟合成。对于点的检测和定量，目前采用的算法可分为高斯拟合(Gaussian fitting)和拉普拉斯算子高斯拟合(Laplacian of Gaussian)两大类[1~4]。除了在单块胶上获得点的信息外，在一系列不同胶上监测某个蛋白质点的变化是生物学分析所感兴趣的问题。不同图像间的比较首先要确定足够的坐标(landmark)，即每块胶上都存在的点。这些坐标用于确定图像之间几何转换的多项式，然后用于其他点的匹配。表达量变化的蛋白质点或已经鉴定的点可在软件中进行数据归类和解释，还可以建立数据库通过因特网来进行成果共享。

双向图谱分析软件的开发始于 20 世纪 70 年代末，至今经历了三代的发展历程。最初的程序是使用 DEC PDP11 家族的小型计算机在 VMS 操作系统上运行，如 Elsie，LIPS，Gellab 和 Tycho；第二代是基于 Unix 的程序，盛行于 80 年代后期，如 Gellab-Ⅱ，Elsie-4，Kepler，Melanie 和 Quest；第三代是基于 Unix、NT 和 Mac 多种计算机平台的程序，如 2D BioImage，Melanie-Ⅱ和 Phoretix 2D。当今，双向图谱分析软件正向 PC 化发展，而且界面更友好，使用更方便，功能更强大，以适应这一工具进入普通实验室，如 PDQuest，Melanie-Ⅲ，Phoretix 2D 和 ImageMaster 2D 等。近年来随着蛋白质组学的兴起，新的双向凝胶电泳分析软件在自动化方面又有较大改进，如 Z3、Delta 2D 和 Progenesis。

第二节 PDQuest 软件简介

PDQuest 是显示(imaging)、分析(analyzing)双向凝胶电泳图谱和数据库查询(databasing)的一个软件包，由 Bio-Rad Laboratories 开发，目前已经升级到第 7 版。它运行在 Windows 环境下，图形界面简单易用，有标准菜单、工具条，并且支持键盘操作。PDQuest 从 Bio-Rad 的专业扫描仪中获取图像，用高级算法去除图像背景噪声、凝胶中的人为假象和水平垂直的条纹，然后用点细化(spot segmentation)的方法去检测和定量蛋白

质斑点，最后将结果显示并能输出到其他的统计分析用途中。它能同时分析凝胶上的10 000个蛋白质斑点，能跟踪报告所有蛋白质形式和样品浓度的变化，并给出合理的定量和定性数据。PDQuest 的匹配(matching)和数据(databasing)分析能力使它能同时分析几百块凝胶，柱形图(histograms)方式能比较一组实验中每块凝胶上相同点的量的变化。使用者能将相关的蛋白质斑点归为一类，再从实验胶上产生的参考图像(master image)上做高级分析和比较。扫描的文件能被插入或输出到发现组(discovery series)中，文件格式转换为 TIFF 格式以便作其他用途。图像匹配的数据能输入到列表中(spreadsheets)以便做简单的分析。

用 PDQuest 分析 2D 凝胶包括以下主要步骤。

①看图和解释。看图工具可对图像进行缩放、翻转、任意角度调整、切割和去除背景噪声，还可将图像转换为数据。另外，利用密度工具还可看到整张图上点的 OD 值的变化。在点的检测中 PDQuest 自动标记哪些点是过饱和点(saturated)，哪些是模糊点(faint)，这些点将不能参加以后的分析过程。使用者也可人工标记哪些点不参与分析过程，还可将紧挨在一起的点综合成一个点，便于比较和分析。

②点的检测和定量。PDQuest 点的检测过程是自动进行的，但是使用者可人工设定参数，如扫描过程、看图过程、点检测的灵敏度以及是否纠正模糊点等。PDQuest 还可根据所做的标准给出凝胶上每个点的相对分子质量(mr)和等电点(pI)信息。

③凝胶匹配。根据不同的目的做出不同的匹配组(matchset)，然后将它们整合成标准(standard)。选择标准胶的要素有：点的数目、代表胶、对照胶和点的质量。对匹配组(matchset)可进行编辑操作：部分匹配、手工匹配、检查非稳定点(erratic spot)、重新匹配、添加点等。匹配过程中需要人工设定一些参考点(landmark)。

④数据归类。PDQuest 可将凝胶上的点做数据处理，有 6 组设置可使用：定量、定性、任意(arbitrary)、布尔(boolean)、统计和解释(annotation)。

⑤报告输出。匹配组(matchset)的数据可输出到列表中进行分析，扫描的图像可以 TIFF 的形式输出到其他图像处理软件中进行分析。

图谱分析是一个多步骤、有顺序的过程，包括蛋白质斑点的检测、高斯拟合(Gaussian fitting)、定量、背景扣除、图像定位(alignment)，最后是凝胶图像的比较，蛋白质斑点配比(matching)。图谱分析的大多步骤是自动进行的，但由于凝胶电泳的质量，通常有 10%左右的斑点没有被正确检测，有小部分蛋白质斑点可能会错配，需要人工校正。要提高蛋白质斑点检测和配比的正确率，一方面需要提高双向凝胶电泳的质量，另一方面需建立更为精确的斑点检测和配比的计算机算法。最近，Panek 等[5]提出了一种称为 point pattern matching 的新的计算机算法，检测蛋白质斑点的正确率大于 95%，斑点配比的精确率大于 98%。

第三节　PDQuest 软件的使用

一、材　　料

1. 硬件

最低硬件配置要求为 Pentium 166 处理器，64M 内存，3G 硬盘，17″显示器，1024×

768 分辨率,256 色,SCSI 接口。Windows 95/98/NT 操作系统。

我们实验室的硬件配置为 Pentium Ⅱ 350 处理器,256M 内存,15G 硬盘。Windows ME 操作系统。

2. 软件

PDQuest 双向凝胶图像分析软件 6.2.0 版, Bio-Rad Laboratories, Hercules, CA。

二、方 法

(一) 图像采集

染色完毕的凝胶用水洗净后,置于 Bio-Rad GS710 扫描仪上,胶上盖一块干净的滤光板,注意胶和扫描仪以及滤光板之间不能有气泡。

扫描仪可以从文件菜单或者从快速向导中选择。扫描仪面板上的"选项"里有扫描仪校准的设置,软件安装后要根据扫描仪随机附送的校准表来重新设置。扫描仪会根据此表在每次扫描前自动将机器校准一遍。

扫描的第一步是选择所要扫描的介质的种类,如凝胶、X 射线片等,然后选择染色方法,如银染、考马斯亮蓝染色等。软件会根据所选择的介质和染色方法来改变发光方式和滤光片的设置。第二步,预扫描,确定所要收集的数据的范围。第三步,数据采集。此时会提示给扫描文件命名。扫描仪面板见图 15.1。

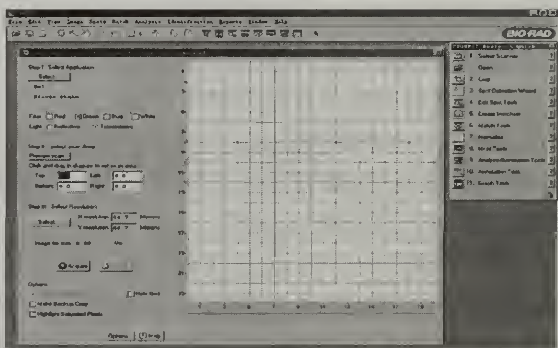


图 15.1 扫描仪面板

(二) 看图和编辑

在 View 菜单中 Zoom In、Zoom Out 可以对图像的进行缩放,View Density 工具可以看到胶上点的光密度变化。Advance View 工具提供了更多的看图方式,在点的检测和编辑中可以用到。Windows 菜单下的 Tile 可以平铺图像,Imitate 用来使所有其他胶的显示和所选胶的显示相同,比如,用 Zoom out 观看一块胶的某一放大区域时,用 Imitate 可以很方便地看到其他胶上此放大的区域。Assign 和 Interchange 工具在建立点的检测以后会用到,用于在指定窗口内显示同一块胶所产生的不同文件,如 2D scan 文件、Image File 文件和 Gel Spot 文件。看图菜单和它们的图标见图 15.2。

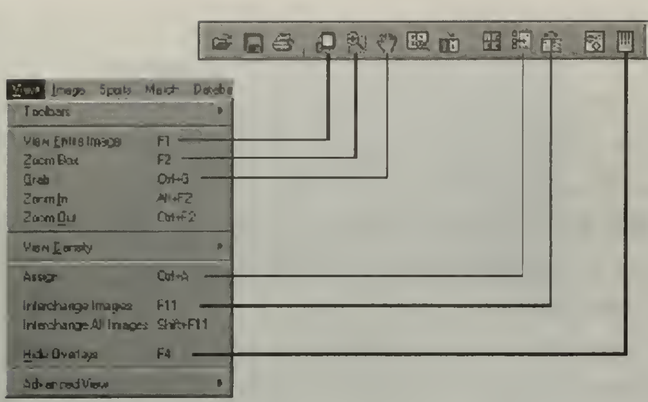


图 15.2　看图菜单和对应的图标

图像编辑时要将所有图像的大小、方向、调整一致，同时还要去除图像上的背景噪声。

打开图像后用 Image 菜单下的 Rotate 或 Flip 工具将所有图像的方向调整一致，然后用 Filter 工具去除图像中的背景噪声。背景噪声分为两类，称为 Pepper 和 Salt。Salt 的光密度比背景浅，Pepper 的光密度比背景深。它们都属于 Outlier 类型的噪声。噪声有三种密度分布方式，称为 Gaussian、Uniform、Outlier，用 View 菜单下的 View Density 可以判断噪声的分布方式。然后选择过滤的像素大小，像素越大过滤的效果越明显。需注意，过滤对图像来说是不可逆的操作，因此，最好在保存的时候选择备份(copy and filter)。过滤的提示框见图 15.3。

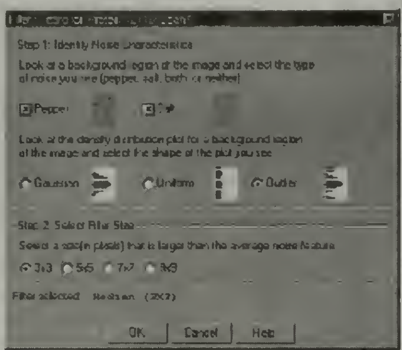


图 15.3　过滤的提示框

去除噪声后，用 Image 菜单下的 Crop 命令可切除胶上四周一些不需要的部分。高级的切割方法是在所有图像中选择一块胶做参考胶，用 Image 菜单下的 Advance Crop 工具中的 Define crop area 定义切割区域，然后执行 Place Crosshair 命令在胶上选择一个分离得好、又很清晰的点做为十字准线(crosshair)。Save Crop Setting 保存此切割设置。其他胶只要 Load Crop Setting，找到十字准线后执行 Crop 命令即可。这样能保证切割出来的胶规格一致。

Image 菜单中的 Transform 可以随意调整图像的亮度和对比度，以便于看图，但它并不改变图像内在的光密度。图像的亮度和对比度可以手工调整，也可以选择自动方式(Auto-Scale)。

（三）点的检测和编辑

PDQuest 6.1 版增加了点检测的参数设置向导(Spot Detection Parameter Wizard)。打开胶图像后，选择一块分离效果比较好的胶作为标准胶，然后在 Spot 菜单下选择 Spot Detection Parameter Wizard，会弹出如图 15.4 所示的界面。

点检测时需要设定以下参数：最模糊的点、最小的点、最大的点和胶的背景。Spot Dectection Parameter Wizard 面板左边的一列图标是看图工具，可以放大、缩小、移动图像，并能看到胶上点的光密度变化，这些工具是为了便于找到最模糊的点、最小的点和最大的点。最模糊的点即是图中可见的光密度最低的点，它确定了点检测时的灵敏度。最小的点为面积最小的点，它确定了点检测的尺寸范围。在大多数情况下，最模糊的点即是最小的点，因此可以略去设定最小的点这一步。最大的点为面积最大的点，它确定了去除背景时的背景抽提转球和条纹去除圆盘的半径。胶的背景设定应选择胶上一块尽可能大的、没有点和条纹的区域作为胶的背景，胶的背景确定了点检测过程中的噪声过滤方式。



图 15.4　点检测的参数设置向导

点的参数设置好后，按 Find Spot Centers 测试一下点的检测效果。软件将会给每个检测得到的点打上十字标记。如选中 Gaussian Model During Test，则检测到的点将会减少，因为聚焦在一起的点可能将被认为是一个大的点。要获得满意的点的检测效果，点的检测需要重复若干遍，直至找到最佳的参数。

在 Spot Dectection Parameter Wizard 面板上的 Optional Controls 选项里，有 Streaks、Background、Speckles 和 Options 4 栏。在点的参数设置好后，Background 栏和 Speckles 栏一般不用改动，如果条纹的去除不够多的话，可将 Streaks 中的半径降低一半再重新检测一遍。

背景抽提有 3 种方法：Floating Ball、Roller Ball 和 Bi-disk。默认的 Floating Ball 方法和球半径是最佳的。如果默认方法对背景的抽提太多或有图像假象的话，可以选择 Roller Ball 方法，但它通常会产生更多的图像假象。Bi-disk 方法比 Floating 和 Roller Ball 方法稍快一点，对均匀光滑的背景处理效果比较好，对不均匀光滑的背景处理效果较差。

背景抽提的方法见图 15.5。

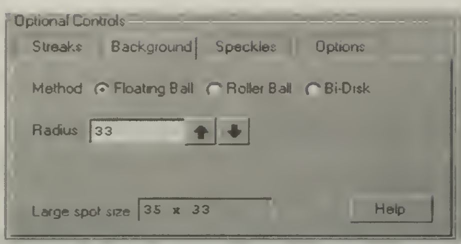


图 15.5　背景抽提的方法

Speckles 栏的设置用于过滤背景噪声，Perform Smoothing 一般要选中，过滤的类型是根据噪声的种类来选择的，见表 15.1。

表 15.1　噪声的种类

噪声类型	过滤类型
Pepper, Salt Noise	Median
Gaussian Noise	Weighted Mean
Salt, Gaussian Noise	Power Mean
Pepper, Gaussian Noise	ContraMean

Kernel Size 的选择在最小点和噪声的平均像素大小之间，像素的大小可用放大看图工具来观察。见图 15.6。

Options 栏的 Fix Faint Spots 是默认选中的，Faint Spots 是指一些光密度低于一定值(如 0.2 OD)又高于背景的点，Fix Faint Spots 将这些点的 X、Y Sigma 值赋为同一 Y 轴座标的点的值。它解决了像素非常小的点的定量问题。见图 15.7。

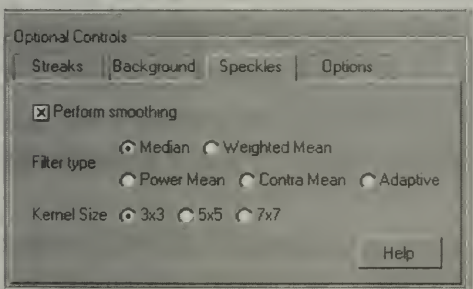


图 15.6　Kernel Size 的选择

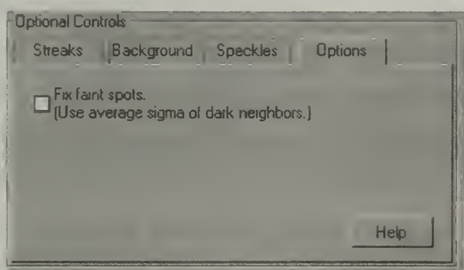


图 15.7　Fix Faint Spots 的选择

点的检测参数设置好后，在 Parameter Set 栏按 Save 将此设置保存，Load 此设置可对其他胶进行点的检测。Processing All Gels 可选择设置参数检测多块胶的点。

点的检测完成后加上原来的 2D scan 文件，另外又生成两个文件：Filtered 和 Gaussian。同时打开这三个文件，利用 Spot 菜单的 Add Spot 和 Remove Spot 可增加或去除点，在 More Edit Tools 里还可将几个重叠在一起的点合并成一个点，或者手工标记哪些点是

饱和点,哪些是模糊点。Spot Boudary Tools 是 6.2 版新增的功能,如果一个点是形状不规则的饱和点,或者点的质量很差,那么在点的检测过程中 Gaussian Modeling 不能精确合成这个点,用 Spot Boundary Tool 可手工描绘这些点,但点的定量可能不很准。

(四) 建立和编辑比较组(Matchset)

建立和编辑比较组的菜单见图 15.8。

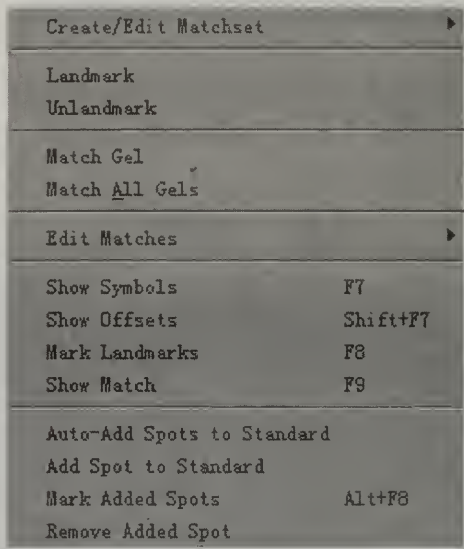


图 15.8　建立和编辑比较组的菜单

点的检测完成后,同一个胶文件会自动生成 3 个文件:Gel Spot、Gel Image 和 2D Scan。建立比较组,命令如下:Matchset 菜单下 Creat Matchset。然后选择要比较的文件,Gel Spot 和 Gel Image 文件也应包含在内,这样在建立比较组后,还可以做点的编辑。按 OK 键之后会出现一个对话框,要选择一块胶做标准胶,标准胶最终会包含所有比较组成员的点,所以,标准胶应选择一块最具代表性的胶,比如实验中的对照,而且这块胶要分离得好。标准胶也可在比较组建立后再选。比较组分为两个水平:低水平和高水平。低水平比较组建立在原始的扫描胶文件上,而高水平的比较组则产生于低水平比较组的标准胶上。

点的编辑最好在 Matchset 建立之后,这样可以同时编辑多块胶,也便于看到各块成员胶上的情况。

各块胶之间比较,首先要确定几个坐标(Landmark),Landmark 的选择标准如下:①所有胶上都有;②分离得好,分辨率高;③靠近图像的四角;④独立的没有其他点与其重叠;⑤位于或靠近点密集的地方。胶的各个部分应至少选择 5 个点做为 Landmark,一般应标记约 10%的点做为 Landmark。

Landmark 建立之后,软件会自动匹配所有其他点。用 Show Symbols 可以看到匹配的结果,绿色表示匹配的点,红色表示该点只在该块胶上有,而其他胶上没有。Show Offsets 可以看到点匹配的情况,如果出现以下结果则表示匹配可能有问题:①Offset 有交叉;

②某一区域 Offset Line 不平行;③Offset Line 长短不均;⑤某块区域只有极少的 Offset,暗示这个部分尚未匹配。

Erratic Spots 用于显示在各个成员胶上有量的不同的点,量的差别可以自由设定显示。如输入 3,则成员胶上有 3 倍量的差别的点,将会出现在标准胶上,以黄色显示。

Manual Match 用于在自动匹配情况下未匹配的点,用手工方式匹配这些点。匹配完成后,一些只在成员胶上有而标准胶上没有的点需要加到标准胶上。Add Spot To Standard 会完成这一任务。一般情况下,从成员胶上加到标准胶上的点不会很多。Colony Match 是针对同一批样品而言,比较只需要定义两个 Landmark,就可以快速比较两块胶。

(五) 比较组的数据分析

Analysis 的菜单见图 15.9。



图 15.9 Analysis 的菜单

一般在建立和编辑好比较组之后,需要分析各个成员胶上蛋白质点差异表达的情况。

Spot Review Tool 可以看到各个点的定量柱型图,并且方便地观察所有点在各块胶上的定量变化。

Image Stack Tool 将多幅图像进行堆栈,然后逐幅观察,可方便地观察多幅图像上某一斑点的有或无的表达情况。

Scatter Plot Tool 可分析比较成员胶间所有点的相关系数。

成员胶之间的非差异表达的量的变化可以通过 Normalize 来平衡。非差异表达的点的量的变化,可能由以下因素造成:

①样品制备中所用的细胞数目不同;②不同试剂和实验方法而导致样品制备的效率不同;③制备中的操作误差导致样品丢失;④样品制备过程中移液枪的误差;⑤上样时样品丢失;⑥两块胶的染色时间不同;⑦放射性标记的效率不同;⑧不同荧光试剂的激发/释放效率不同;⑨图像扫描时的检测能量不同;⑩图像扫描时胶的暴露时间不同。

Normalize 的选项如图 15.10 所示。

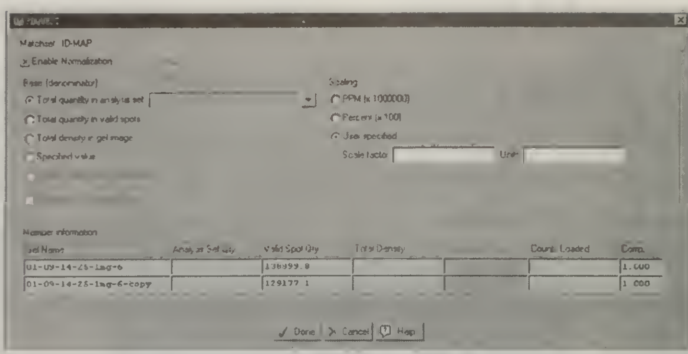


图 15.10 Normalize 的选项

Normalize 方法取决于实验要求和实验方法，如仅仅想鉴定一些胶上持家蛋白质(housekeeping protein)，则应对 Total Spots In Analysis Set 进行 Normalize，如果已知了某些蛋白质点的量的变化，则选择 Specified Value，输入已知的变化因素。

Normalize 的公式：Normalized Spot Quantity＝(Raw Spot Quantity * Scaling Factor * Pippeting Compensation Factor)/Normalize Factor。

Raw Spot Quantity 是未校准的点的量；Scaling Factor 是选择的中间值，一般是 10^6；Normalize factor 是由 Normalize 的方法决定的。

Normalize 的方法如下所述。

①对一套分析组中总的点。如持家蛋白质，即所有 Matchset 成员都有的点，但这首先要求从 Standard 上做一个 Arbitrary Analysis Set，其中包含了持家蛋白质，选择多个持家蛋白质的好处是极个别的点被平均掉了。

②对 Total Of All Valid Spots。All Valid Spots 是指所有标准胶上的点。在这个方法中，成员胶上每个点的原始定量值除以该块胶上所有被包含在标准胶上的点的定量值。这个方法假设实验中只有极少点的变化，而且这些点在胶上是平均分布的。在不知道样品变化的原因时，这种方法是有用的。

③对 Total Density In Gel Image。这个方法假设胶与胶之间的总密度(点的密度加上背景密度)不变。在不知道样品变化的原因时，这种方法是有用的。

④对特定值。如果已知样品的变化，如实验中的上样量和总细胞数，以及图像检测时的时间长度，可以针对特定值来设定。

⑤对 Counts Loaded From Gel Record。Counts Loaded Into Gel 是在做校准时在 Gel Record 里设定的值。这种方法只有在 Matchset 胶里做了校准时才是可行的。Counts Loaded Into Gels 列在 Normalization Dialog 表的末尾。这种方法假设实验中只有极少的点的变化，而且这些点在胶上是平均分布的。如果已知样品量的变化，正如在扫描时设定的，这种方法是有用的。这种方法选择 Pippetting Error Compensation，可以补偿所有比较组成员间的加样误差。当这个窗口被选中后，软件会试图纠正微小的系统误差，如扫描计数、加样等的误差。校准是假设这些误差对所有蛋白质的影响是均等的，同时假设实验中如果有影响时只有极少数蛋白质会发生变化，而且它并不改变从一个样品到下个样品间的随机分布。当此窗口选中时，一个对话框会显示每个比较组成员的计算的误差值。这

些值会出现在 Normalize Dialog 表的尾端。Pippetting Error Compensation 能纠正由于上样量的不同而引起的系统误差，但它存在两个假设前提：①log 点是正常分布的；②点的变化是随机的，如果细胞在加药状态下，那么它的合成不是随机的。补偿因子（compensation factor）用于计算平均比率，如果假设的第二点成立的话，此值应为 1。

同一实验条件下、一批胶中由于染色等效果的不同会造成部分点定量不准。Replicate Group 会平均这些实验中的误差。Replicate Group Quantitation 会显示平均点的定量信息、变异系数和 Replicate Group 中的胶的数目。

双向凝胶电泳可以用商品化的 2D standard 做为内标，用来计算其他胶上蛋白质的等电点和相对分子质量。用 pI Standard 工具可输入一些标准蛋白质的相对分子质量和等电点。然后软件会自动标出其他点的相对分子质量和等电点。但相对分子质量和等电点超过标准蛋白质范围的点，其相对分子质量和等电点将无法显示。

（六）点的计数和解释

点的计数和显示在 Analysis 菜单下的 Spot Numbers 里。每个匹配的点都将分配一个标准点数 SSP（standard spot number）。除了 SSP 之外，还有一个计数系统 DSN（database spot number），DSN 系统的好处是它根据相对分子质量、等电点画出的格子暗示了每个点的定位。

（1）SSP

SSP 会在比较组建立之后自动分配，SSP 由 4 位数组成，前两位数代表该点所在部分的 X、Y 轴，后两位数代表它在方框里的计数。每个方框里的数目不超过 100 个。如果超过 100，那么软件会从其他方框里调用数目。

（2）DSN

DSN 在输入内标的等电点和相对分子质量之后，从 Database 菜单下 Spot Number 中的 Assign DSN Grid，可绘出其他点的 DSN 数目。DSN 也由 4 位数组成：ABnn，A 代表相对分子质量，B 代表等电点，nn 是方框里的数目。

第四节　双向凝胶电泳软件的实际应用

一、双向凝胶电泳的重复性和定量定性差异分析

比较组成员间的蛋白质差异表达分析和双向凝胶电泳的重复性分析，需要结合 PDQuest 和 Office 软件包里的 Excel 来进行统计学分析。

（一）双向凝胶电泳的重复性分析

双向凝胶电泳的重复性关系到不同实验室间数据交换和利用，也直接影响蛋白质组研究的价值和应用。分析的过程如下：不同批次的胶建立好比较组之后，在 File 菜单下的 Export 里输出比较组里所有匹配点的信息（X、Y 轴坐标，点的定量）。一般是随机挑选 10%左右的点，在 Excel 里算出这些点的 X、Y 轴坐标的标准差，作为双向凝胶电泳的位置重复性指标。定量的重复性用这些点的定量变异系数作为指标。

（二）差异表达分析

差异表达可用 Creat Analysis Sets 分析，Creat Analysis Sets 菜单如图 15.11 所示。

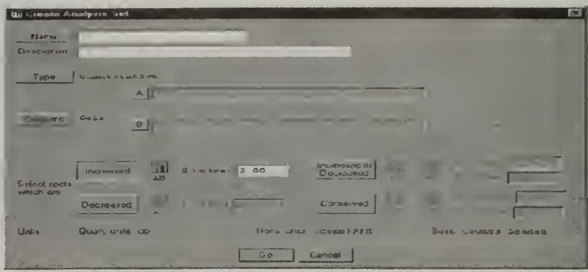


图 15.11　Creat Analysis Sets 菜单

Set 的选择类型有 4 种，见图 15.12。

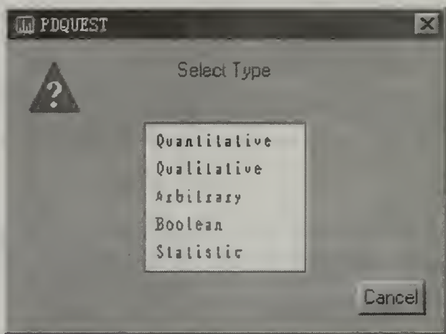


图 15.12　选择类型

其中 Quantitative 用于定量（多/少）差异表达分析，差异表达的倍数可自由设定。Qualitative 用于定性（有/无）差异表达分析。Arbitrary 可用于自由显示某些点，如某块胶特有的点、匹配的点和自己设定的点。Boolean 用于分析多重 Set 里点的或、与等分析。Statistic 是对 Replicate Group 里的点做统计学分析，分析有 4 种 Student T-Test（$P<0.05$）、Log scaled Student's T-Test、Mann-Whitney Test。常用的差异表达分析选用 Student T-Test。

我们对正常肝细胞系 L02 和肝癌细胞系 BEL7404 进行了双向凝胶电泳图谱的重复性分析[5]，结果见表 15.2。

表 15.2　双向凝胶电泳图谱的重复性分析

参数	位置重复性		定量重复性	
	等电聚焦方向	SDS-PAGE 方向	BEL-7404	L-02
标准差（mm），mean	0.73±0.38	0.44±0.23		
标准差(mm)，range	0.09～1.74	0.08～1.24		
变异系数(%)，mean			19.2±10.9	17.6±9.8
变异系数(%)，range			3.7～59.9	1.9～58.8

定量差异表达分析，在 BEL7404 细胞系里高表达的蛋白质点如表 15.3 所示。

表 15.3　在 BEL7404 细胞系里高表达的蛋白质点的定量差异表达分析

蛋白质索引[a]	等电点[b]	分子质量[b]/kDa	相对体积[c]/%		P 值
			BEL-7404	L-02	
154	—	—	0.051±0.003	0.017±0.004	0.0000
155	—	—	0.125±0.022	0.066±0.019	0.0067
232	—	—	0.030±0.005	0.016±0.003	0.0042
436	5.96	64.3	0.064±0.012	0.034±0.002	0.0030
453	—	—	0.025±0.004	0.013±0.005	0.0081
464	—	—	0.395±0.041	0.307±0.020	0.0085
522	—	—	0.127±0.011	0.105±0.005	0.0091
591	6.46	57.1	0.189±0.015	0.107±0.010	0.0001
601	6.55	56.8	0.068±0.006	0.032±0.005	0.0001
636	6.11	57.4	0.071±0.014	0.038±0.009	0.0061
675	—	—	0.082±0.007	0.035±0.014	0.0010
693	5.97	55.7	0.017±0.007	0.003±0.002	0.0097
735	5.8	53.7	0.085±0.012	0.038±0.002	0.0002
750	—	—	0.051±0.004	0.020±0.008	0.0003
770	—	—	0.064±0.021	0.007±0.004	0.0019
784	—	—	0.034±0.014	0.005±0.002	0.0057
812	—	—	0.257±0.021	0.203±0.018	0.0076
813	—	—	0.259±0.008	0.020±0.015	0.0000
817	—	—	0.069±0.018	0.020±0.009	0.0033
845	6	50.2	0.052±0.010	0.021±0.008	0.0025
846	—	—	0.304±0.065	0.074±0.040	0.0010
852	6.16	49.9	0.164±0.019	0.123±0.004	0.0066
855	—	—	0.401±0.135	0.126±0.009	0.0066
858	5.76	49.8	0.049±0.010	0.015±0.003	0.0006
916	—	—	0.219±0.014	0.096±0.014	0.0000
1002	5.4	44.8	0.028±0.007	0.009±0.005	0.0059
1043	—	—	0.126±0.007	0.045±0.008	0.0000

续表

蛋白质索引[a]	等电点[b]	分子质量[b]/kDa	相对体积[c]/%		P 值
			BEL-7404	L-02	
1098	—	—	0.100±0.007	0.057±0.018	0.0038
1184	6.92	35.0	0.018±0.005	0.006±0.004	0.0072
1189	—	—	0.131±0.030	0.011±0.006	0.0002
1256	—	—	0.073±0.006	0.051±0.006	0.0022
1287	—	—	0.041±0.009	0.015±0.007	0.0037
1298	—	—	0.054±0.012	0.016±0.001	0.0008
1303	5.91	31.6	0.156±0.024	0.070±0.009	0.0005
1325	—	—	0.109±0.016	0.067±0.011	0.0049
1347	5.38	31.0	0.044±0.009	0.013±0.008	0.0022
1390	5.11	30.1	0.131±0.019	0.063±0.017	0.0017
1397	5.32	29.9	0.180±0.037	0.100±0.016	0.0074
1410	—	—	0.141±0.010	0.050±0.016	0.0001
1425	—	—	0.110±0.012	0.074±0.014	0.0073
1446	6.64	27.2	0.041±0.009	0.019±0.007	0.0091
1480	5.32	27.1	0.481±0.058	0.025±0.013	0.0000
1521	6.74	25.0	0.099±0.008	0.057±0.010	0.0006
1540	5.23	24.7	0.127±0.020	0.075±0.003	0.0025
1554	5.38	23.6	0.017±0.002	0.009±0.002	0.0029
1566	—	—	0.062±0.010	0.026±0.007	0.0009
1568	5.42	22.2	0.100±0.009	0.049±0.008	0.0001
1589	—	—	0.119±0.022	0.041±0.007	0.0006
1606	—	—	0.094±0.023	0.045±0.008	0.0074
1635	—	—	0.233±0.026	0.123±0.020	0.0005
1665	—	—	0.223±0.014	0.156±0.027	0.0044
1732	—	—	0.466±0.091	0.085±0.069	0.0005

a. 参考胶中的索引；b. "—"表示没有对分子质量和等电点校准；c. 每个点的体积以其占胶上所有点的体积比表示。

在 L02 细胞系里高表达的蛋白质点如表 15.4 所示。

表 15.4　在 L02 细胞系里高表达的蛋白质点的定量差异表达分析

蛋白质索引[a]	等电点[b]	分子质量[b]/kDa	相对体积[c]/%		P 值
			BEL-7404	L02	
648	—	—	0.173±0.055	0.320±0.037	0.0043
837	—	—	0.169±0.031	0.247±0.022	0.0065
866	6.34	49.2	0.129±0.030	0.240±0.022	0.0010
892	5.3	47.7	0.078±0.080	0.252±0.044	0.0089
894	5.36	47.7	0.010±0.008	0.033±0.004	0.0017
938	5.63	48.5	0.067±0.006	0.118±0.022	0.0043
976	5.52	47.5	0.012±0.005	0.172±0.017	0.0000
1019	5.26	43.9	0.064±0.007	0.090±0.005	0.0010
1025	—	—	0.018±0.015	0.097±0.039	0.0094
1032	5.85	45.2	0.097±0.023	0.148±0.014	0.0092
1071	5.38	42.2	0.036±0.003	0.078±0.010	0.0002
1072	6.35	37.2	0.207±0.073	0.613±0.103	0.0007
1090	—	—	0.091±0.019	0.244±0.061	0.0031
1107	5.38	41.1	0.220±0.011	0.290±0.017	0.0005
1122	6.34	36.5	0.033±0.007	0.107±0.019	0.0003
1237	—	—	0.090±0.007	0.126±0.002	0.0001
1314	—	—	0.045±0.008	0.086±0.020	0.0090
1317	5.64	31.1	0.100±0.026	0.223±0.020	0.0003
1365	5.47	30.7	0.080±0.010	0.167±0.022	0.0004
1406	5.98	29.5	0.171±0.021	0.259±0.026	0.0018
1505	5.29	25.9	0.243±0.050	0.361±0.021	0.0049
1529	—	—	0.050±0.009	0.107±0.005	0.0000
1647	—	—	0.036±0.002	0.052±0.007	0.0052
1688	—	—	0.037±0.011	0.080±0.010	0.0012
1700	—	—	0.065±0.013	0.117±0.021	0.0056
1705	—	—	0.855±0.118	1.600±0.178	0.0004

a. 参考胶中的索引；b. “—”表示没有对分子质量和等电点校准；c. 每个点的体积以其占胶上所有点的体积比表示。

关于定性差异表达分析，只在 BEL7404 或 L02 细胞系里表达的蛋白质点如表 15.5 所示。

表 15.5 在 BEL7404 或 L02 细胞系里表达的蛋白质点的定性差异表达分析

BEL-7404				L02			
蛋白质索引[a]	等电点	分子质量[b] /kDa	相对体积[c] /%	蛋白质索引[a]	等电点[b]	分子质量 /kDa	相对体积[c] /%
2530	6.12	55.6	0.025±0.010	230	—	—	0.015±0.005
2542	6.12	49.1	0.008±0.003	512	5.96	61.2	0.172±0.037
2551	6.09	47.0	0.046±0.014	514	5.87	61.3	0.014±0.006
2555	5.57	46.3	0.016±0.010	904	6.11	49.8	0.010±0.006
2570	5.40	37.2	0.018±0.005	915	5.1	47.3	0.068±0.027
2573	5.18	36.0	0.005±0.001	1050	—	—	0.169±0.058
2584	5.77	28.7	0.005±0.006	1055	—	—	0.285±0.050
				1120	5.57	42.1	0.057±0.014
				1328	6.59	30.9	0.008±0.008
				1359	—	—	0.041±0.015
				1543	—	—	0.004±0.002
				1657	—	—	0.177±0.027
				1673	—	—	0.145±0.044
				1685	—	—	0.767±0.214

a.参考胶中的索引;b.“—”表示没有对分子质量和等电点校准;c.每个点的体积以其占胶上所有点的体积比表示。

以 IPG 胶代替传统的基于两性电解质的等电聚焦,较大程度提高了双向凝胶电泳的重复性。Corbett 等[7]采用 pH4～8,18cm 长的 IPG 胶,12%T 的 SDS 均一浓度胶双向凝胶电泳,同一实验室内 IEF 方向平均位置偏差的最好结果是 0.62mm, SDS-PAGE 方向为 0.44mm;实验室间 IEF 方向偏差为 1～1.5mm,SDS-PAGE 方向为 1～2.6mm,量的变异未做比较。Blomberg 等[8]也做过 IPG-DALT 双向凝胶电泳实验室间的重复性,他们采用线性 pH4～7,13%T,18.5cm 高的 SDS 均一胶,一块胶中分辨蛋白质斑点数最多的达 1 561 点,最少的只有 1 054 点。同一实验室内 IEF 方向位置偏差的最好结果为 0.68%, SDS-PAGE 方向为 0.50%, 量的变异为 19.9%;实验室间 IEF 方向偏差为 2.8mm,SDS-PAGE 方向为 1.8mm,量的变异为 34.5%。我们的实验结果表明[9],IEF 方向位置偏差为 0.73mm, SDS-PAGE 方向为 0.44mm, 量的变异为 17.6%～19.2%。可见双向凝胶电泳中蛋白质斑点位置的重复性较好,但量的重复性不尽如意,尤其是实验室间量的比较。

二、双向凝胶电泳矢量图

蛋白质组研究的一个重要内容是蛋白质各种翻译后修饰的研究。运用图像分析软件,在双向凝胶电泳数据库的参考图上,蛋白质位置如果与该蛋白质的理论预测位置不同,在这两个位置之间连一直线,即为该蛋白质的“矢量”,这样的图谱称双向凝胶电泳矢

量图("Vector maps"of the 2-D gels)[10]。在矢量图上,有的蛋白质具有一个矢量,有的具两个或多个矢量,表明可能存在多种翻译后修饰。如果对不同翻译后修饰与蛋白质双向凝胶电泳迁移位置效应关系进行大量研究,可能建立一个矢量库,从而有助于蛋白质翻译后修饰种类和程度等预测,为实验研究提供信息。

第五节　双向凝胶电泳数据库和网上比较

一、双向凝胶电泳网络数据库资源

蛋白质组的研究目前还正处于展开阶段,双向凝胶电泳作为目前研究细胞总蛋白质最简单、最有效的手段,纷纷被采用,其数据库亦纷纷建立。最著名的 SWISS-2DPAGE 已更新到第 6 版,内容包括 59 条记录,人的 23 张参考图(脑脊液,结肠上皮细胞,结肠腺癌细胞系 DL－1,红白血病细胞系,HepG2,HepG2 分泌蛋白,肝,肾,淋巴瘤,大噬菌体样细胞系 U937,血浆,血小板,早幼粒白血病细胞 HL60 和红细胞),鼠的样品(附睾脂肪垫,腓肠肌肉,肝,胰岛细胞),大肠杆菌(3 个 pH 范围 3.5～10,4～5,4.5～5.5),酿酒酵母和网柄菌属的 discoideum。在这些图上一共鉴定了 2 761 个蛋白质点,用凝胶匹配的占 52.3%,免疫印迹的占 24.7%,微量测序法的占 17.9%,氨基酸组成分析的占 6.5%,共迁移的占 3.2%,质谱占 12.9%。因为许多蛋白质点采用多种方法鉴定,所以累加的百分数要超过 100。SWISS-2DPAGE 数据库可通过关键词(蛋白质名称,副本,作者,全文)或图像点击的方法来查询。我们也可将任何 SWISS-2DPAGE 数据库上查出的蛋白质序列或使用者输入的序列来将它定位到任何一张 SWISS-2DPAGE 的参考图上。

除 SWISS-2DPAGE 外其他的 2-DE 数据库见表 15.6～表 15.12。

表 15.6　多物种的 2-DE 数据库

数据库	研究对象
SIENA-2DPAGE(2D-PAGE database at Department of Molecular Biology, University of Siena, Italy) http://www.bio-mol.unisi.it/2d/2d.html	人乳房导管癌 衣原体 *Trachomatis*,克氏病线虫 *C. elegans*
2-DE database at Max Planck Institute for Infection Biology, Berlin, Germany http://www.mpiib-berlin.mpg.de/	人 Jurkat T 细胞 分枝杆菌 Bovis、Tuberculotis,幽门螺杆菌 *Helicobacter pylori*,疏螺旋体属 *Garinii*
Argonne protein mapping group server http://www.anl.gov/BIO/PMG/	人乳房上皮细胞,鼠肝细胞 火球菌属 *Furiosus*
Research Institute for Biological Science, Science University. Tokyo http://www.rs.noda.sut.ac.jp/～kamom/2de/2d.html	小鼠:小脑,皮层, 海马,回纹 拟南芥的愈合组织,叶,种,茎 水稻叶愈合组织,叶,茎,胚,种子
Large Scale Biology Corp http://www.lsbc.com/	人,小鼠,大鼠肝 玉米,小麦

表 15.7 哺乳动物的 2-DE 数据库

数据库	研究对象
Human and mouse 2-D PAGE databases - Danish Centre for Human Genome Research http://biobase.dk/cgi-bin/celis	人角质化细胞，转变的癌细胞，膀胱多磷细胞癌，尿，纤维原细胞 小鼠肾细胞
RAT HEART-2DPAGE (German Heart Institute Berlin) http://gelmatching.inf.fu-berlin.de/~pleiss/	大鼠心脏细胞
HSC-2DPAGE(Heart Science Centre, HarefieldHospital) http://www.harefield.nthames.nhs.uk/nhli/protein/	人心脏(心房、心室)，内皮细胞 鼠、狗心脏(心室)
HEART-2DPAGE (German Heart Institute Berlin) http://www.chemie.fu-berlin.deberlin.de/user/pleiss/	人(心房、心室)
HP-2DPAGE (Heart 2-DE Database, MDC, Berlin) http://www.mdc-berlin.de/~emu/heart/	人(心室)
PDD Protein Disease Database (NIMH - NCI) http://www-lecb.ncifcrf.gov/PDD/	人血浆，脑脊液，尿
PPDB - PhosphoProtein DataBase http://www-lecb.ncifcrf.gov/phosphoDB/	鼠淋巴细胞
Washington University Inner Ear Protein Database http://oto.wustl.edu/thc/innerear2d.htm	人的外淋巴 豚鼠内淋巴(内耳膜状迷路内的淋巴液)，皮质、耳石、顶盖膜器官
PMMA - 2DPAGE at Purkyne Military Medical Academy, Czech http://www.pmma.pmfhk.cz/	人结肠癌
Rat Serum Protein Study Group at Istituto di Scienze Farmacologiche, Milano, Italy http://users.unimi.it/~ratserum/homeframed.html	大鼠血清(对照 Sprague-Dawley (CD)鼠，松节油鼠)
Rat Serum Protein Database at Aebersold Protein Laboratory, University of Washington, Seattle http://weber.u.washington.edu/~ruedilab/info/ratserum/ratserum.html	大鼠血清(正常、急性相)
BALF 2D Database at Department of Biological Chemistry, University of Mons-Hainaut, Belgium http://www.umh.ac.be/~biochim/BALF2D.html	人的支气管肺泡灌洗液 鼠的支气管肺泡灌洗液
Mito-pick (2-DE gel of mitochondria at CEA, France) http://www.dsv.cea.fr/thema/MitoPick/Mito2D.html	人线粒体
Molecular An atomy Laboratory at Indiana University, Columbus http://iupucbio1.iupui.edu/frankw/molan.htm	人血浆，肝 鼠血浆 大鼠肝，肾(皮层，髓)，脑，血浆，睾丸
Human Colon Carcinoma Protein Database at JPSL, Ludwig Institute for Cancer Research, Melbourne, Australia http://www.ludwig.edu.au/jpsl/jpslhome.html	人胎盘，结肠细胞系
Cambridge 2D PAGE http://sunspot.bioc.cam.ac.uk/NEURON.html	鼠神经元

表 15.8 果蝇的 2-DE 数据库

Drosophila melanogaster, Drosophila melanogaster at the Karolinska Institute	http://try.cmb.ki.se/

表 15.9 酵母的 2-DE 数据库

YPD Yeast Protein Database (registration is required) at Proteome, Inc.	http://www. proteome. com/databases/YPD/
YPM - Yeast Protein Map at Institut de Biochimie et Génétique Cellulaires, Bordeaux, France	http://www.ibgc.u-bordeaux2.fr/YPM
LBYPM- 2D Map of The Kronenbourg Breweries Yeast Proteins at TEPRAL Laboratory, Strasbourg, France	http://www.sdv.fr/tepral/Page1.htm

表 15.10 植物的 2-DE 数据库

Maize Genome Database at INRA, Gif-sur-Yvette, France	http://moulon.moulon.inra.fr/imgd/
2-D gels for maritime pine at INRA, Bordeaux, France	http://www.pierroton.inra.fr/genetics/2D/
Plant Plasma Membrane Database (PPMdb) at Gent University, Belgium	http://sphinx.rug.ac.be:8080/

表 15.11 细菌和病毒的 2-DE 数据库

PHCI-2DPAGE - Parasite host cell interaction 2D-PAGE database, at Department of Medical Microbioloy and Immunology, University of Aarhus, Denmark	http://www.gram.au.dk/
Sub2D - 2D PROTEIN INDEX of Bacillus subtilis, at Department of BioSciences, University of Greifswald, Germany	http://microbio2.biologie.uni-greifswald.de:8880/sub2d/pub/sub2d.ask
Cyano2Dbase - Synechocystis sp. PCC6803, at the Kazuka DNA Research Institute, Japan	http://www.kazusa.or.jp/cyano/cyano2D/
2-D PAGE Aberdeen (Haemophilus influenzae and Neisseria meningitidis), at Department of Medical Microbiology, University of Aberdeen, Scotland	http://www. abdn. ac. uk/~ mmb023/2dhome.htm
SSI-2DPAGE - Mycobacterium tuberculosis, at Department of Tuberculosis Immunology, Statens Serum Institute, Denmark	http://www.ssi.dk/en/forskning/tbimmun/tbhjemme.html
ECO2DBASE- NCBI repository	http://pcsf.brcf.med.umich.edu/eco2dbase/ (new) ftp://ncbi.nlm.nih.gov/repository/ECO2DBASE/ ECO2DBASE (old)

表 15.12 细胞系的 2-DE 数据库

TMIG-2DPAGE - Age-related Proteome Database at Tokyo Metropolitan Institute of Gerontology, Japan	http://proteome.tmig.or.jp/2D/
Embryonal Stem Cells at University of Edinburgh	http://www.ed.ac.uk/~nh/2DPAGE.html
UCSF 2D PAGE (A375 cell line)	http://rafael.ucsf.edu/2DPAGEhome.html
Human leukemia cell lines Lab. de Biochimie et Tech. des Proteines, Bobigny	http://www-smbh. univ-paris 13.fr/lbtp/biochemistry/biochimie/bque.htm

二、双向凝胶电泳数据库的构建

双向凝胶电泳数据库可用 Make2ddb 来建立，该软件包在 Solaris 和 Irix 操作系统中运行，可在 http://www.expasy.ch/ch2d/make2ddb.html 上下载到。它要求至少一个 2-DE 参考图(Melanie 解释的图或 TIFF 图)和普通文本描述(绘图过程，作者，日期等)，Make2ddb 也允许管理和出版网上 2-DE 数据，并且综合了查询链接。

操作过程如下所述。

①安装 make2ddb，安装中必须指定服务器的 bin，cgi-bin 和 htdocs 目录和存放有全部 2-D 数据文件的 2d 目录。

②做参考胶。参考胶上至少要有一组登记号码的标记。

③做文本文件。给出参考胶上每一个鉴定蛋白质的记录——db.dat，并以 SWISS-2DPAGE 的形式保存。

④给 db.dat 加上标题。每条记录用//分开，每个文件加上三行：DB-NAME Release NN，MONTH YEAR，空行。

⑤做 GIF 图。用 make2db 的 melch2dtogif 和 mastergif 程序给每一个参考胶做五个 GIF 图放在 db 目录下。

a) melch2dtogif-x-e-11-z0.1 Master NULL>Master.gif：将 Master.gif 放在目录 2d/small-gifs/Master 下。

b) mastergif-z1.0-11-e Master；mv Master _ id.gif Master _ id _ big.gif.

c) mastergif-z0.6-11-e Master.

d) Melch2dtogif-e-11-z1.0 Master NULL>Master _ big.gif.

e) Melch2dtogif-e-11-z0.6 Master NULL>Master.gif.

将 4 个 GIF 图放在 2d/gifs/Master 下，这些文件包含了参考图上鉴定和未鉴定的蛋白质点。

⑥做 existing.map 文件。建立 2d/current _ release 的子目录，将包含所有参考图名称的 existing.map 文件放在其下。

⑦修改主页。将 make2db 提供的主页 2d.html 稍做修改以适合自己的需要并放在 htdocs/2d 目录下。

⑧给 2d 加个化名。编辑 srm.conf 文件给 htdocs/2d 加个化名：Alias /2d xxx/2d/ . xxx是去 htdocs 的目录，这将保证链结时找到正确的地址。

⑨运行 Make2ddb。在 2d/current_release 目录下为每条记录建立独立的文件,为关键词和作者查询建立索引,在 cgi-bin 目录下存放版本信息。

⑩通过互联网测试双向电脉数据库[11]。

三、双向电脉凝胶的网上比较

随着互联网上双向电脉数据库的增多,实验人员越来越容易在网上找到与自己样品类似的双向电脉数据库,从而将实验结果与之进行比较。比较时要用到 Flicker 程序,这是用 Java 语言编写的程序,可在任何装有 IE3.0 或 Netscape2.0.1 以上版本的计算机上运行。该程序可在 http://www-lecb.ncifcrf.gov/flicker/flicker.zip 上下载到。

运行 Flicker 时终端使用者和数据库服务器的关系如图 15.13 所示。

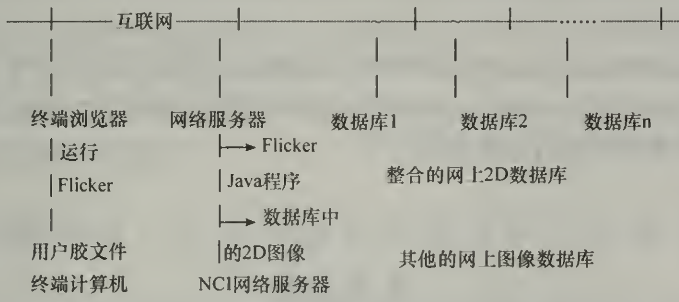


图 15.13　运行 Flicker 时终端使用者和数据库服务器的关系

用户登录到 NCI Flicker 网页 http://www-lecb.ncifcrf.gov/flicker 上,Flicker 程序可将终端用户的双向电脉凝胶图像与远程数据库的图像数据进行实时比较,比较的结果将在第三个窗口显示,使用者也可用鼠标将一个图像拖到另一个图像上进行比较。有三种比较方式。

1. 图像闪烁比较(flickering)

因为双向电脉凝胶图像有很多扭曲变形,而且图像上每个部分的变形程度也不相同,所以两块胶不可能一次性匹配得非常好,这就需要在凝胶上不同的部位做一些标记(landmark),然后进行多次比较,被形象地称之为闪烁比较(flickering)。

如果一个点及其周围的区域和另一张图都匹配得非常好,我们可认为这个点与另一张图上已知的点是相同的,如果完全鉴定这个点则还需要序列分析、氨基酸组成分析、质谱、单抗或其他方法。

2. 图像增强效应(enhancement)

为了改善闪烁比较的质量,在比较前可做一些图形处理,如空间弯曲——在不改变灰度值的情况下使一张图的某一区域和另一张图的相应区域在几何上完全对应。另外,还有一些处理方法如图像尖锐(image sharpening)技术——一种边缘增强效应,即在原始灰度的图像上增加陡峭成分或 laplacian 点检测功能和增加图像的对比度。

3. 图像加工变形(image-processing transform)

(1) Affine 空间弯曲变形

变形后的坐标(u,v)与原始坐标(x,y)的关系如下:u(x,y) = ax + by + c; v(x,y) = dx + ey + f

a, b, c, d, e, f 是 Flicker 根据每块胶上的三个标记计算出的值。变形的结果是使两块胶比较的区域在几何上等同。

(2) 假 3-D 变形

假 3-D 变形可增加重叠的点,变形后点的坐标(x′,y′)和原始坐标(x,y)的关系如下:dx= width * sin(q); x′=[dx * (height − y)/height] + x; y′ = y − zScale * g(x, y)。灰度值决定了 z scale 的大小,q 值在 −45° − 45°范围内。

(3) 边缘尖锐效应

边缘尖锐(sharpening)效应可增加模糊点的边缘。处理过的点的坐标'(x, y)与原始点的关系如下:'g(x, y)=[e scale * edge(x, y)+(100 − e scale) * g(x, y)]/100。公式中 e scale 值用来衡量增加的边缘检测值[12]。

(丁士健　俞利荣)

参 考 文 献

1 Anderson NL, Taylor J, Scandora AE *et al*. The TYCHO system for computer analysis of two-dimensional gel electrophoresis patterns. Clin Chem, 1981,27: 1807～1820

2 Appel RD. Interfacing and intergrating databases. In: Wilkins MR ed. Proteome research: new Frontier in funtional genomics. Vol 1, Berlin: Springer-Verlag, 1997, 149～175

3 Lemkin PF, Lipkin LE. GELLAB: a computer system for two-dimensional gel electrophoresis analysis. Ⅲ. Multiple two-dimensional gel analysis. Comput Biomed Res, 1981,14: 272～297

4 Lemkin PF, Lipkin LE. GELLAB: a computer system for 2D gel electrophoresis analysis Ⅱ. Pairing spots. Comput Biomed Res, 1981,14: 355～380

5 Wu Y, Lemkin PF, Upton K. A fast spot detection algorithm for 2D electrophoresis analysis. Electrophoresis, 1993, 14: 1341～1356

6 Panek J, Vohradsky J. Point pattern matching in the analysis of two-dimensional gel electropherograms. Electrophoresis, 1999,20: 3483～3491

7 PDQuest User Guide for Version 6, BIO-RAD

8 Corbett JM, Dunn MJ, PoschA *et al*. Positional reproducibility of protein spots in two-dimensional polyacrylamide gel electrophoresis using immobilised pH gradient isoelectric focusing in the first dimension: an interlaboratory comparison. Electrophoresis, 1994,15: 1205～1211

9 Blomberg A, Blomberg L, Norbeck J *et al*. Interlaboratory reproducibility of yeast protein patterns analyzed by immobilized pH gradient two-dimensional gel electrophoresis. Electrophoresis, 1995,16: 1935～1945

10 俞利荣, 王楠, 吴高德等. 人肝癌细胞系 BEL-7404 和正常肝细胞系 L-02 表达的蛋白质组双向凝胶电泳分析. 科学通报, 2000,45: 170～178

11 Ron D, Appel RD, Hoogland C *et al*. Constructing a 2-D database for the world wide web. 2-D Proteome analysis protocols. In: Link A J, ed. Methods in molecular biology. Vol 112, Totowa: Humana Press, 1999, 411～416

12 Lemin PF. Comparing 2-D electrophoresis gels across internet databases. 2-D Proteome analysis protocols. In: Link AJ ed. Methods in molecular biology. Vol 112, Totowa: Humana Press, 1999, 393～410

第十六章　生物质谱技术和蛋白质鉴定

在蛋白质组学研究流程中，蛋白质鉴定是最关键的一环。如果说在蛋白质组分离技术方面出现各种方法并存的局面，那么，蛋白质鉴定的主流技术的确定已无可争议，那就是质谱技术。

蛋白质组学研究与传统的蛋白质研究最重要的区别在于前者是大规模研究，因此在研究策略上必然是从整体出发的大视野，但落实到蛋白质组中的单个蛋白质成分的研究技术，并没有脱离蛋白质研究的基本框架。可以认为，蛋白质组学大规模技术的发展必然促进了对单个蛋白质的研究，而很多在单个蛋白质研究中得以应用的技术也很快应用到蛋白质组学研究中。正因为两者相辅相成、密不可分的关系，在本章中，作者并不打算将两者割裂，而将对蛋白质和蛋白质组学研究中的质谱方法做一总述，介绍质谱技术在蛋白质分析鉴定中的发展和应用。

第一节　生物质谱基本原理和工作模式

就技术方法而言，蛋白质研究技术比核酸研究技术要相对复杂和困难得多：不仅氨基酸残基种类远多于核苷酸残基，而且蛋白质有着复杂的翻译后修饰如磷酸化，糖基化等，这给分离分析带来很多困难。由于蛋白质结构层次复杂，物理化学特性多样，因此在蛋白质结构功能研究中，技术方法的发展对研究的影响十分显著。如 Edman 降解氨基酸测序使蛋白质一级结构得以阐明；NMR 的出现大大推进了对蛋白质溶液构象的解释。

如果说上述这些技术的应用已成为蛋白质研究的里程碑，那么 20 世纪 80 年代末以来，在生物大分子领域中得以广泛应用的质谱技术则可以视为一个新的里程碑[1~5]。质谱技术在 20 世纪初就已出现，但一直仅应用于有机小分子领域，直到 80 年代才渐渐进入到生物大分子领域。随着 20 多年来的发展和应用，质谱技术已是蛋白质研究中必不可少的工具，并成为蛋白质组研究中的主要支撑技术。

对蛋白质和多肽而言，质谱技术就是要确定一个蛋白质或多肽的分子质量。而由此所获得的信息，涉及的领域已远远超过了检测分子质量这一简单应用。因为分子质量是一个蛋白质、多肽或氨基酸最基本的特征，这种特征是专一的。不同的氨基酸组成、不同的氨基酸序列、不同的修饰方式、不同的蛋白质间结合方式、不同的位点差异都可以在蛋白质的分子质量上体现。对蛋白质以及组成蛋白质的多肽，氨基酸的分子质量测定实际上就是对蛋白质种类和性质的鉴定 (identification)。

质谱技术的基本原理是样品分子离子化后，根据不同离子间质荷比（m/z）的差异来分离并确定分子质量。一台质谱仪一般由进样装置、离子化源、质量分析器、离子检测器和数据分析系统组成。在这几部分中，离子化源和质量分析器是两个中心部件，也是发展得最快的两个方面，不同类型的质谱仪主要就是根据这两个部件来命名的，如电喷雾（ESI）、快原子轰击质谱（FAB）是根据离子化源不同来区分，而飞行时间质谱、四极杆质谱

则是根据质量分析器来划分。不同离子源和质量分析器相匹配,就基本确定了质谱仪的工作方式。如 ESI 离子源通常与四极杆、离子阱质量分析器相匹配,而基质辅助的激光解析离子化源通常与飞行时间飞行器相结合,即所谓 MALDI-TOF 质谱仪。表 16.1 提供质谱特别是生物质谱技术的发展历程。质量分析器的原理在本书的上篇已有详尽的阐述,在此着重讨论生物质谱的离子化方式和相关技术及其应用。

表 16.1 质谱技术发展历史简介

研究者	时间	贡献
J.J Thompson	1912	第一台质谱仪
Gohlke and McLafferty	1956	气相质谱
Biemann, Seibl *et al*.	1959	肽序列分析
McLafferty, Jennings *et al*.	1967	串联质谱
McLafferty	1973	液相质谱
Comisarow and Marshall	1974	傅里叶回旋共振质谱
Barber, Bordoli *et al*.	1981	快原子轰击
Yamashita, Fenn *et al*.	1984	电喷雾离子化用于生物大分子
Tanaka, Karas *et al*.	1988	基质辅助激光解吸离子化用于生物大分子

(一) 生物质谱的软电离方式

早期的质谱技术仅应用于小分子研究领域,其中一个重要的原因是当时采用的电子轰击(electron impact, EI)电离能量太强,而生物大分子结构复杂且热不稳定,EI 不适于生物大分子分析。直到 20 世纪 80 年代,电喷雾离子化和基质辅助激光解吸离子化的出现和应用,才使质谱技术真正进入到生物大分子分析领域,而这两种技术的原创研究者也都获得了 2002 年的诺贝尔化学奖。

1. 电喷雾离子化

电喷雾离子化(electrospray ionization, ESI)是一种"软电离"方式,它是在"离子蒸发"原理基础上发展起来的一种离子化方法[6]。离子蒸发是指离子从液相发射到气相的过程。待测分子溶解在溶剂中,以液相方式通过毛细管到达喷口。在喷口高电压作用下形成带电荷的微滴,随着微滴中的挥发性溶剂蒸发,微滴表面的电荷体密度随半径减少而增加,到达某一临界点时,样品将以离子方式从液滴表面蒸发,进入气相。这一过程即实现了样品的离子化,由于没有直接的外界能量作用于分子,因此对分子结构破坏较少,是一种典型的"软电离"方式。电喷雾离子化示意图见图 16.1。

电喷雾质谱的另一大特点是可形成多电荷离子,因此,在较小的 m/z 范围内可以检测到大分子质量的分子。采用电喷雾质谱目前可测定分子质量 100 000Da 以下的蛋白质,最高达 150 000Da。由于离子蒸发使电喷雾质谱采用液相方式进样,因此可与蛋白质化学中常用的液相色谱联用,即液相色谱-电喷雾质谱(LC-MS)。蛋白质或多肽经过 HPLC 分离后,直接进入质谱进行分子质量测定。

在常规电喷雾中,喷雾时形成的较大液滴中,有一部分样品还来不及离子化,因而降

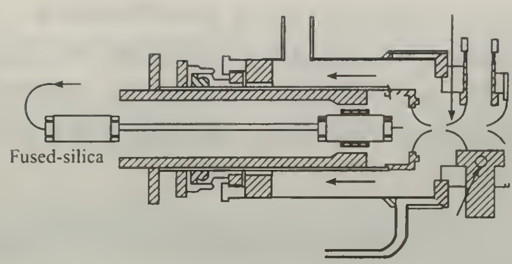


图 16.1 电喷雾离子化示意图

低了样品的利用率和灵敏度。近年发展的纳升电喷雾方法（nanospray），采用内径为 1μm 左右的喷针，样品注入到喷针的喷口，产生喷雾，在喷口前端喷雾液滴体积是常规电喷雾 1/100，使样品被充分利用，并有效离子化[7]。纳升电喷雾以极低的流速（20～40nl/min），消耗极少的样品（1～2μl），灵敏度达低 fmol。

2. 基质辅助激光解吸离子化

基质辅助激光解吸离子化（matrix-assisted laser desorption ionization，MALDI）是将样品均匀包埋在固体基质中，基质吸收激光提供的能量而蒸发，携带部分样品分子进入气相，并将一部分能量传递给样品分子使其离子化。

早期的激光解吸离子化没有基质辅助，激光能量直接作用于被分析物，使分子碎裂，产生大量碎片，而不易得到分子离子。1988 年，Tanaka 等首先采用可吸收激光的基质（如甘油）将被分析物包埋其中，避免了激光能量对分子结构的破坏，测定了分子质量为35 000Da 的蛋白质[8]。后来发现，采用固体基质可以使样品分散得更均匀，离子化效果更好[9,10]。通常采用的激光束波长为 337nm 。由于样品的电离过程是由基质介导的，因此基质的选择对分子离子化有很大影响，继而影响到分析的灵敏度、分辨率和精确度。合适的基质应保证待测物的离子化，而且基质本身的离子造成的背景较弱。对于蛋白质和多肽样品，较为通用的基质有芥子酸（sinapic acid，SA），α-氰基-4-羟肉桂酸（α-cyano-4-hydroxycinnamic acid，CHCA），2,5-羟基苯甲酸（2,5-dihydroxybenzoic acid，DHB）。MALDI 最大的特点是离子电荷通常为 1 或 2 个，而不像 ESI 中为多电荷，对分子质量较大的样品而言，不会形成复杂的多电荷图，因而对图谱的解析比较清楚。图 16.2 为 MALDI 离子化的示意图。

由于 MALDI 的进样是固相方式，不像 ESI 的液相进样容易受溶液性质的影响，因此对杂质的忍耐性较好。通常在进行生物样品制备时，往往会在样品中引入一些缓冲液或

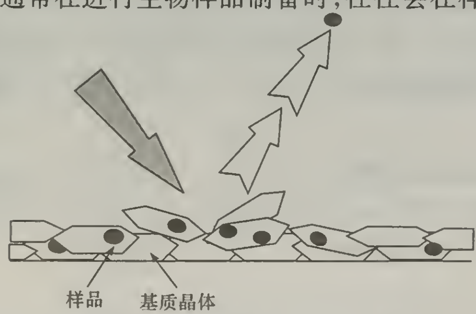


图 16.2 MALDI 离子化的示意图

去垢剂(如尿素和 SDS)等。这时采用 MALDI 会有较好的效果。一般 MALDI 采用的激光是在紫外范围,而近年来发现可用红外激光分析生物大分子,红外激光的能量比紫外激光弱,产生的碎片离子较少,比较合适分析不稳定的样品,如磷酸化肽等[11,12]。也有人提出不使用基质,而将样品直接置于多孔硅板上进行激光解吸离子化,从而大大减少了基质的背景离子[13,14]。

(二) 常用生物质谱的类型和特点

生物质谱的离子化方法基本采用 ESI 和 MALDI,质量分析器则有不同的选择,如三极四极杆、离子阱、飞行时间、傅里叶回旋共振等。不同的质量分析器各有其优势和特点,也有不同的应用范围,目前发展趋势是将不同类型的质量分析器串联起来,以提高质谱的工作性能和适用范围。在此以离子化方式为中心对生物质谱仪进行分类和描述。

1. 电喷雾质谱仪

目前常用的以电喷雾为离子化方式的质谱有:电喷雾-三级四极杆质谱(ESI-TQ-MS)、电喷雾-离子阱质谱(ESI-IT-MS)、电喷雾-三级四极杆-飞行时间质谱(ESI-Q-TOF-MS)、电喷雾-傅里叶回旋共振质谱(ESI-FTICR-MS)等。

2. 基质辅助激光解吸离子化质谱仪

目前已有基质辅助激光解吸离子化-飞行时间质谱(MALDI-TOF-MS,MALDI-TOF/TOF-MS)、基质辅助激光解吸离子化-三级四极杆-飞行时间质谱(MALDI-Q-TOF-MS),基质辅助激光解吸离子化-离子阱质谱(MALDI-IT-MS)等。

表 16.2 和表 16.3 为离子化方式和各种质谱性能的比较。

表 16.2 ESI 和 MALDI 的比较

电喷雾方式	可在线联用的方式	离子电荷形式	对杂质容忍程度
ESI	液相	多电荷	直接进样时差,与液相色谱联用后好
MALDI	固相	一般为 1 或 2 个电荷	较好

表 16.3 各种类型质谱的性能比较

质谱类型	MS/MS 功能	速度	分辨率	灵敏度	精确度
电喷雾-三级四极杆质谱(ESI-TQ-MS)	+ +	+ +	+ +	+ + +	+ +
电喷雾-离子阱质谱(ESI-IT-MS)	+ + + +	+ +	+ +	+ + +	+ +
电喷雾-三级四极杆-飞行时间质谱(ESI-Q-TOF-MS)	+ +	+	+ + +	+ + +	+ + +
电喷雾-傅里叶回旋共振质谱(ESI-FTICR-MS)	+ + + +	+	+ + + +	+ + + +	+ + + +
基质辅助激光解吸离子化-飞行时间质谱(MALDI-TOF-MS)	+	+ + + +	+ +	+ + + +	+ +

续表

质谱类型	MS/MS 功能	速度	分辨率	灵敏度	精确度
基质辅助激光解吸离子化-串联飞行时间质谱(MALDI-TOF/TOF-MS)	++	+++	++	++++	++
基质辅助激光解吸离子化-三级四极杆-飞行时间质谱(MALDI-Q-TOF-MS)	++	+++	+++	+++	+++
基质辅助激光解吸离子化-离子阱质谱(MALDI-IT-MS)	++++	+++	+	++++	+

注:+ 表示性能强大程度。

(三) 生物质谱与其他技术的联用方式

液相色谱与电喷雾质谱(LC-ESI-MS)连用,可对色谱分离的成分直接用质谱进行在线(online)分析,而不需要收集这些成分[15,16]。由于蛋白质传统分析中采用的肽谱(peptide mapping)分析是在液相色谱上进行的,因此 LC-ESI-MS 的出现使肽谱分析不再仅仅根据色谱中的保留时间(retention time)来区分多肽,而是采用多肽分子质量来对其进行鉴定,大大提高了分析的效率和精确度。在灵敏度方面,由于色谱的灵敏度与色谱柱的内径相关,内径越小,灵敏度越高。肽谱分析传统上采用内径为 2.1mm 或 1.0mm 的柱子,连接的电喷雾喷头内径较大,这是最初的 LC-ESI-MS 模式,称为 normal-spray,灵敏度在低 pmol 水平。后来采用的毛细管液相色谱(capillary LC),柱内径可在 100~300μm 之间,流速为 1~15μl/min[17~19]。相应的电喷雾喷头内径也缩小了,称为 microspray,这种 cLC-ESI-MS 的灵敏度可达 fmol。进一步又发展了 100μm 以下柱内径的纳升流速(nanoflow)的色谱[20,21],与纳升喷雾头(nanospray)联用,灵敏度可达低 fmol。液相色谱的分离方式也有多种选择,除了分析肽谱常使用的反相柱外,还可采用分子筛、亲和柱、离子交换柱等。无论何种液相色谱与质谱联用,都要求色谱洗脱液挥发性较好,以保证电喷雾离子化的效率。

(四) 串联质谱与碰撞诱导解离

串联质谱(tandem MS,MS/MS)是指多个质量分析器相连,分离母离子,进行碰撞解离,并检测子离子[22,23]。MS/MS 较早在四级杆质谱中实现方法是:将三个四级杆串联,第一个四级杆进行母离子分析,选择感兴趣的离子进入第二个四级杆,与惰性气体碰撞成碎片后,进入第三个四级杆进行子离子分析 (图 16.3)。离子阱和傅里叶转换离子回旋共

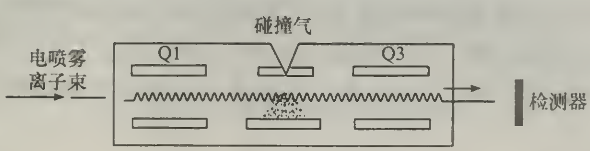


图 16.3 串联质谱(MS/MS)示意图

振质谱(FTICR-MS)也可以进行 MS/MS[24,25]。选择的母离子阱内与惰性气体碰撞而产生碎片离子。与三级四极杆通过空间分离方式获得母离子和子离子不同,离子阱和傅里叶共振离子回旋质谱采用的是时间分离方式,因而可以进行多级 MS (MS^n),可达 10 级(MS^{10})。FTICR-MS 还出现双阱设计,分别进行离子储存和离子分析,更加强了多级 MS 功能[26]。与串联质谱平行的一个概念是碰撞诱导解离(collision induced dissociation, CID)[23,27]。在没有出现串联质谱时,为获得分子结构信息,需要在离子源内(in-source)对离子进行碰撞,使其碎裂。源内 CID 灵敏度高,但由于没有选择性,因此碎片的专一性不强。串联质谱出现后,逐渐取代了源内 CID。严格地说,CID 仅指离子解离成碎片的过程,串联质谱则包括了母离子选择、CID 和子离子分析三个过程。另外,在电喷雾质谱中,离子从离子源输送到质量分析器过程中,要经过一些孔道,如 orifice 或 skimmer,有时可以通过改变这些孔道上的电压,使离子发生一定程度的解离。这一方法可以用于分析翻译后修饰蛋白质和非共价结合蛋白质。

第二节　质谱法分析完整蛋白质和多肽

对完整蛋白质或多肽分子质量的确定,是蛋白质结构分析的重要部分。对于点突变的蛋白质、不均一的糖蛋白、多位点修饰的磷酸化蛋白质、N 端封闭或残基被共价修饰的蛋白质而言,质谱法是快速、准确的鉴定方法。将测定的分子质量与理论分子质量比较,可以验证或推测蛋白质被修饰的情况,或是否有突变位点;还可根据不同亚型在质谱中的相对强度,分析蛋白质被修饰的程度。对于溶液内纯度较高的单个蛋白质或多肽,若采用电喷雾质谱,由于形成了多电荷离子,分子质量越大,图谱越复杂,一般电喷雾质谱可测定的分子质量范围低于 70 000Da,尤其以 3 000 Da 以下的测定结果较为理想,而 MALDI-TOF-MS 由于主要形成单电荷,图谱简单,可测定的分子质量范围可达 100 000Da,甚至 1 000 000Da 以上。

胶内蛋白质的精确分子质量是鉴定蛋白质的指标之一。可将蛋白质转移到 PVDF 膜上,用 MALD-TOF-MS[28]分析,还有人将整块双向凝胶电泳的胶转 PVDF 膜上,整张膜与 MALD-TOF-MS 的基质反应,再进行紫外或红外 MALD-TOF-MS 扫描,分析蛋白质的分子质量[29,30]。甚至有人用质谱代替双向凝胶电泳中的第二向(SDS-PAGE),直接扫描测定等电聚焦后的凝胶,测定蛋白质的精确分子质量,并称之为"virtual 2-D gel"[31]。也可直接将蛋白质从胶内电洗脱或有机溶剂洗脱到液相后,用 MALD-TOF-MS 和 ESI-MS 测定分子质量[32~34]。但大分子或疏水蛋白质抽提到液相较困难。而且,蛋白质经过凝胶分离后,某些氨基酸可能会被丙烯酰胺修饰,或被氧化,也可能被抽提的有机溶剂所修饰(如甲酸),造成分子质量偏移,或造成同种蛋白质以不同形式存在[35]。就目前的大多数质谱精确度而言,根据电泳所得到的等电点和质谱测定的分子质量,仍不足以鉴定蛋白质。以大肠杆菌为例,等电点为 5.5,分子质量为 25 000Da 的蛋白质就有几十种[36]。因此,还需要采用质谱分析蛋白质酶解肽谱的方法来鉴定蛋白质。

第三节 质谱法对蛋白质和多肽一级结构的分析及鉴定

一、质谱分析多肽序列的原理

蛋白质测序法是分析蛋白质和多肽一级结构的传统方法。但蛋白质测序法只能鉴定末端(N端或C端),要获得蛋白质所有氨基酸的序列,还需将蛋白质进行酶解,酶解肽段经色谱分离收集后,分别测序,再拼接成全序列。这个过程十分繁琐。而质谱方法可以一次分析蛋白质酶解的所有肽段,解析序列,从而获得蛋白质完整的一级结构。而对于分子质量较小的多肽,只需根据分子质量和MS/MS数据,即可分析其一级结构。

质谱法测定多肽序列有几种方式:首先是串联质谱法或CID法,即在质谱中获得肽段的碎片[23,37]。肽段在质谱中的碎裂有一定的规律,通常最容易断裂的是肽键,形成的是N端离子和C端离子,分别命名为b离子和y离子,其他N端离子命名为a离子和b离子,其他C端离子为x离子和z离子,图16.4为肽链在质谱中断裂形成的主要离子。除了这些主要离子外,侧链基团的断裂也产生一些离子,如Thr,Ser容易脱水,Arg易脱胍。此外,两个以上肽键同时断裂将产生一些内部离子(internal ions)。通常根据碎片离子进行肽序列分析,主要依据由肽键断裂而产生的离子,即b离子和y离子。质谱法测定肽序列,最擅长的是确证已知序列和分析已知序列中的翻译后修饰。而对未知序列的分析,目前还是一个难题。

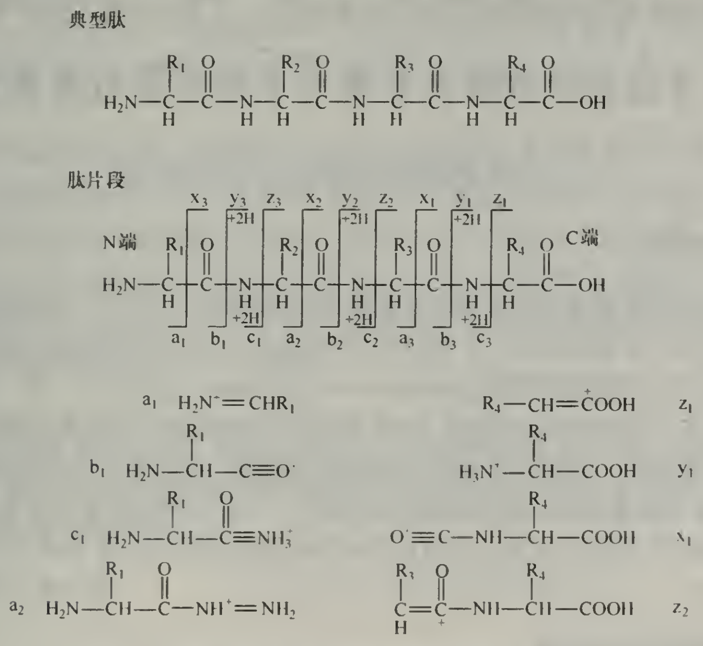


图16.4 肽在CID中的断裂

具有源后衰减(post source decay, PSD)功能的MALDI-TOF-MS也能对肽链测序[38,39],但由于在这一过程中肽链断裂的得率并不高,且产生的碎片离子准确度不够,在

图谱解析时存在一定困难，因此使用的效果不甚理想[40]。近两年出现了将 MALDI 源与其他质量分析器如离子阱、Qq-TOF、TOF-TOF 相连的模式[40~42]，这些分析器都有真正的 MS/MS 功能，因此实现了将 MALDI 的高通量与串联质谱的序列分析功能的合二为一。

由于 Lys 和 Gln 的分子质量很相近（相差 0.0364Da），在分辨率很高的质谱如FTICR-MS 或高分辨磁质谱中，可以区别 Lys 和 Gln 微小的质量差异[43,44]。但在常规的质谱中不易区分。可用多级 MS 断裂 Lys 和 Gln 不同的侧链来区分两者[45]，或将 Lys 衍生为高精氨酸[46]。而 Ile 和 Leu 的分子质量相同，显然不能仅凭一级 MS 分辨，Ile 和 Leu 在 CID 和 CAD 作用下，也会产生不同的侧链离子[47]。最近的研究发现可将多肽与游离铜离子形成复合物，在进行 CID 时 Ile 和 Leu也会产生不同的碎片[48,49]。

另一种测定肽序列的质谱法方法是酶法阶梯式测序[50,51]（ladder sequencing）。一般多应用于 C 端序列分析，即用氨肽酶或羧肽酶依次切除肽链或蛋白质的 C 端氨基酸，测定酶解前后的肽分子质量，通过质谱差异确定何种氨基酸被切割，依此类推末端序列。这一方法多用于 C 端序列测定。但由于酶解是一个动态过程，各个氨基酸酶解速度不同，并且受相邻氨基酸的影响，因此这一方法的通用性不强。目前阶梯式测序一般可测定3～10 个氨基酸。还有报道用质谱法直接测定由化学降解的氨基酸，分析蛋白质末端序列[52~54]。

采用质谱法分析蛋白质的质量肽谱，获得每一肽段的分子质量和序列后，即可对蛋白质进行鉴定。这个方法可用于验证由 cDNA 推出的蛋白质序列，而且特别适用于鉴定基因工程重组蛋白质是否与预期的序列符合，是否均一，突变位点是否准确等。

二、质谱法结合数据库检索对多肽和蛋白质进行鉴定

（一）肽质量指纹图谱法鉴定蛋白质

1. 原理和特点

这一方法的理论基础是认为每个蛋白质经过酶解成为长短不一的肽段后，同一时间获得所有肽段分子质量而形成一个肽段分子质量图谱，这一图谱对蛋白质应是专一而特异的，因此称为肽质量指纹图谱（peptide mass fingerprinting，PMF）。这一方法不需对图谱进行人工解析，只需将实验获得的 PMF 与蛋白质数据库中蛋白质的理论 PMF 比对，就可以鉴定该蛋白质。因此比传统方法速度快、通量高，是最早用于大规模蛋白质鉴定的质谱方法，也是目前最简便的蛋白质鉴定方法之一。近年来随着蛋白质数据库或由基因组数据库和 EST 库等衍生的蛋白质数据库的快速增长和完善，使 PMF 方法的适用性大大增强。

2. 影响结果的关键因素

(1) *灵敏度*

无论如何，能否灵敏地获得 PMF 信号是任何后续分析的基础。灵敏度的欠缺会造成信噪比高，一定程度上又影响到精确度。另外，可检测到的肽段少，对数据库检索是不

利的。

(2)精确度

显而易见,PMF方法主要是靠对肽段的分子质量测定来获得准确的检索结果,因此分子质量精确度是最重要的指标。目前已有延迟萃取和反射检测的MALDI-TOF-MS,以及新近出现的MALDI-Q-TOF-MS,分子质量精确度都可达5ppm左右,已经完全符合数据库检索的要求。但是在实际工作中,尤其是微量样品,或从胶内酶解获得的样品,精确度往往不够理想,在很大程度上会影响检索结果,误差越大,检索准确性越低。

(3)质量校正

鉴于精确度对PMF结果的决定性作用,质量校正就是实验中必不可少的一步,尤其对TOF质量分析器,质量校正非常重要而且需常规进行。质量校正有外标和内标校正两种,外标校正是指在样品测试前用已知精确分子质量的标准品对仪器进行校正。不同的分子质量范围有不同的标准品,通常做PMF时会采用5～10个标准肽的混合物,内标校正是指将已知精确分子质量的标准品掺入到样品中,或将样品中已知精确分子质量的信号为标准对仪器进行校正,显然内标校正比外标校正要精确,不过标准品掺入的步骤较繁琐,标准品浓度若太高,反而会将样品信号抑制,但如果控制得好,对PMF的精确度还是很有帮助的。在对蛋白质的胰蛋白酶酶解产物进行PMF分析时,有时会将胰蛋白酶的自酶解峰作为内标;如果使用牛胰蛋白酶,一般在MALDI-TOF-MS上最明显的自酶解峰理论上为2 163.057,可作为校正标准,但仅能做单点校正,仍不够理想。

(4) 分辨率

分辨率在一定程度上与精确度相关,尤其在肽段分子质量在2 000Da以上时,在MALDI-TOF-MS中的分辨率会有所下降,导致峰变宽,精确度也下降。

(5)数据库

由于PMF鉴定蛋白质是依赖数据库的,因此数据库本身是否完善和准确也是关键因素之一。实际工作中经常会出现PMF图谱很好,但查不到结果的情况。原因之一就是数据库不全,所研究的蛋白质在数据库中不存在,即可能是未知蛋白质。目前已有的蛋白质数据库中,以部分微生物尤其是原核微生物和酵母的数据库最完善,在大部分真核生物中,即使基因组序列已知,由于有内含子、翻译后修饰等因素影响,蛋白质数据库仍不完善,是制约质谱法鉴定蛋白质的一个瓶颈。

(6)样品制备

影响PMF法鉴定蛋白质结果的人为因素主要是样品制备。虽然MALDI质谱对杂质的容忍度较好,但样品制备的好坏会直接影响到图谱的灵敏度、精确度和分辨率,具体的步骤和注意事项见相关章节。

3. 检索方法和工具

(1) 现有的检索工具[55,56]

如何对PMF的实验结果和理论图谱进行比较和评价,是将实验数据转换成具有生物学意义的结果的关键。目前已发展了不少算法和工具,这些软件都可以在互联网上免费使用。互联网上的软件使用的都是公用数据库,如果要使用自主数据库则需购买in-house版。常用的软件及其介绍如表16.4。

表 16.4　PMF 常用的软件

软件名称	网　址
Mascot	www.matrixscience.com
Profound	Prowl.rockefeller.edu
PeptIdent	www.expacy:ch/tools
ProteinProspector	Prospector.ucsf.edu
PepSea	195.41.108.38

(2)Mascot 使用方法

图 16.5 为 Mascot 查询页面,参数的选择如下。

Database:选择数据库,一般采用 NCBI 的非冗余库(NCBInr)。

Taxonomy:根据样品来源选择物种名称(如 Human 数据库)。

Enzyme:选择所使用的酶(如 Trypsin)。

Allow up to:选择可能遗漏的酶切位点数。

Fixed modifications:选择固定修饰。

Variable modifications:选择可变修饰。

Protein mass:选择蛋白质分子质量上限。

Peptide tol. ±:选择肽分子质量误差范围。

Mass values:用以搜索的数值是分子离子质量或分子质量。

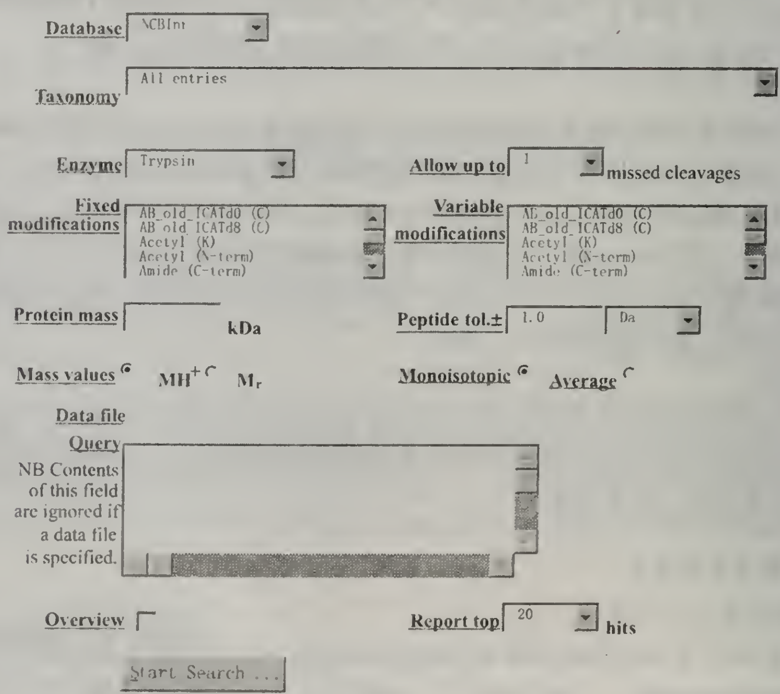


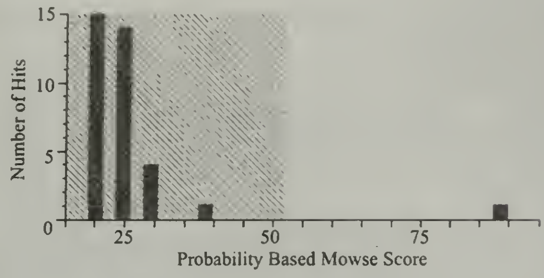
图 16.5　Mascot 查询页面

Monoisotopic:是单同位素分子质量还是平均分子质量。

Data file: 选定文件。

Query: 贴入用于查库的分子质量。

图 16.6 为 Mascot 检索结果的示例。



1. gil6319314 **Mass**: 69896 **Total score**: 89 **Peptides matched**: 11

(NC_001133) Stress-seventy subfamily A; Ssa1p [Saccharomyces cerevisiae]

2. gil6323196 **Mass**: 17433 **Total score**: 38 **Peptides matched**: 4

(NC_001144) Homology to rat S27a; Rps31p [Saccharomyces cerevisiae]

图 16.6 为 Mascot 检索结果

图中的横坐标为鉴定的蛋白质的可能性分数,分数越高,代表结果越可靠,第一位结果和第二位结果分数相差越大,第一位结果越可靠。

(二)一级质谱(MS)结合串联质谱(MS/MS)法鉴定蛋白质

1. 原理和特点

由于 MS/MS 具有结构分析功能,可以获得检测到的每一肽段的序列。因此,根据一级质谱(MS)测得的精确分子质量和串联质谱(MS/MS)所得的序列信息,通过蛋白质数据库的检索,无疑大大增加了蛋白质鉴定的准确性。这一方法多在具有 CID 功能的质谱中进行,MALDI-TOF-MS 的 PSD 功能虽然也能获得一些序列信息,但效率不高,操作繁琐,故应用有限。

2. 影响结果的关键因素

(1)灵敏度、精确度和分辨率

由于有 MS/MS 信息,对图谱的精确度和分辨率的要求不如 PMF 高。因此,影响结果的最重要指标是灵敏度,尤其是在液相色谱-质谱联用的仪器上。通常会采用纳升喷雾提高检测的灵敏度。

(2)MS/MS 图谱质量

MS/MS 图谱的质量是最影响检索结果的因素,除了灵敏度、精确度等指标外,图谱的峰型分布也很重要。理论上,如果所有 b 离子和 y 离子都被检测到,并有同等离子强度时,检索的匹配最理想,但在实际的图谱中,某些离子可能会被遗漏或强度很低,造成匹配程度不理想。

3. 检索方法和工具

(1)现有的检索工具[57,58]

常用的串联质谱检索工具如表 16.5 所示，这些检索工具有的可从网上获得，有的绑定在质谱仪软件中可供使用。我们实验室常用的是 SEQUEST 软件，它与 Finnigan 公司的质谱仪整合，因而可以直接识别原始图谱，而在网上使用的软件则需从原始图谱中抽取峰信号后使用。

表 16.5　常用的串联质谱检索工具

软件名称	网　址
SEQUEST	Fields. scripps. edu/sequest
Mascot	www. matrixscience. com
PepFrag	Prowl. rockefeller. edu
ProteinProspector	Prospector. ucsf. edu

(2) SEQUEST 使用方法

①先在 Bioworks 软件中打开一个原始图谱(Raw Data)。

②在 Action 栏点击 TurboSEQUEST Search，即出现检索参数(图 16.7)。

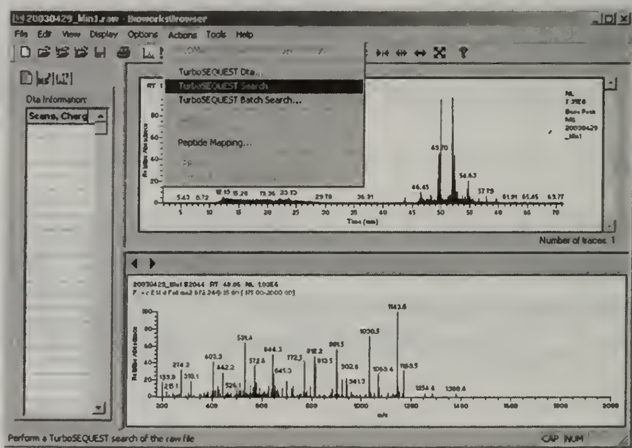


图 16.7　SEQUEST 检索界面

③选择 Basic 参数，见图 16.8。

Scan limits：选择用于检索的扫描数。

MW Range：选择用于检索的肽分子质量范围。

Threshold：选择用于检索的信号下限。

Database：选择数据库。

Options：选择数据库的性质。

Enzyme：选择使用的蛋白酶。

Number of internal cleavage sites：选择可能遗漏的酶切位点数。

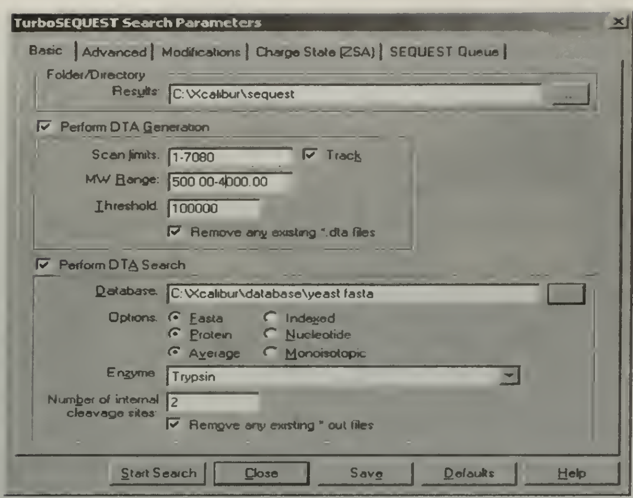


图 16.8　Basic 参数选择

④选择 Advanced 参数(图 16.9),一般可使用默认值。可根据需要选择的参数有:

Precursor Mass:母离子分子质量差异在此范围内可视为同一种离子,可合并检索。

Minimum Group Count:合并检索组的最小值。

Minimum Ion Count:用于检索的最少离子个数。

Peptide:用于检索的肽分子质量误差。

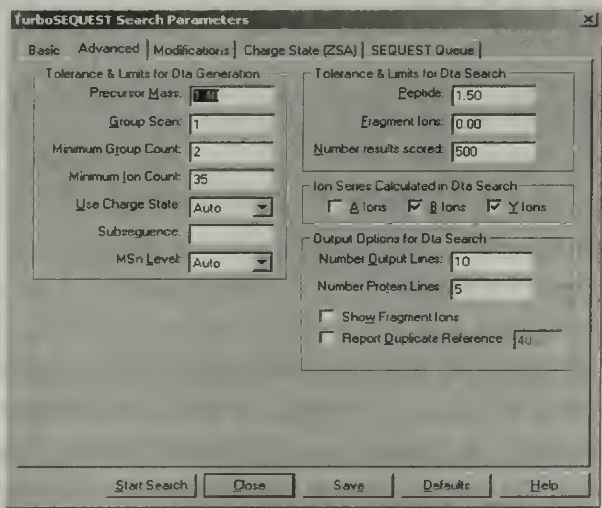


图 16.9　Advanced 参数选择

⑤选择 Modification 参数(图 16.10)

Static:选择固定修饰。

Differential:选择可变修饰。

在 SEQUEST 软件中,通过实验数据与理论数据的相关性(cross correlation)分析就可以鉴定蛋白质。表 16.6 为一个胶内蛋白质的部分数据库鉴定结果。

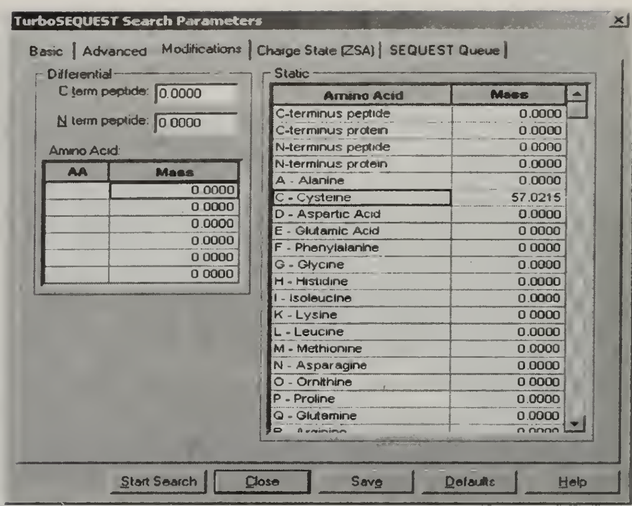


图 16.10　Modification 参数选择

表 16.6　蛋白质的部分数据库鉴定结果示例

File	MassI	MassA	Xcorr	DelCn	Sp	RSp	Ions	Reference	Sequence
2D-14.0974.0989.2	1543.9	1543.8	5.0044	0.535	2032.8	1	21/28	TPIS HUMAN	(R)KQSLGELIGTLNAAK
2D-14.0992.0995.2	1416.1	1415.6	2.3091	0.261	553.0	1	14/26	TPIS HUMAN	(K)QSLGELIGTLNAAK
2D-14.1007.1010.2	1516.3	1513.7	1.1531	0.041	90.8	65	7/24	SPCA HUMAN	(R)DGLLARRDALREK
2D-14.1022.1037.2	2192.8	2193.4	5.6444	0.657	1167.4	1	24/38	TPIS HUMAN	(K)VPADTEVV C * APPTAYIDFAR
2D-14.1040.1049.2	1808.7	1809.1	6.1178	0.609	3068.3	1	27/34	TPIS HUMAN	(K)VAHALAEGL GVIAC * IGEK
2D-14.1055.1088.2	1603.0	1603.9	4.5333	0.540	1501.7	1	20/28	TPIS HUMAN	(K)VVLAYEPVWAIGTGK
1. TPIS HUMAN				152 (14,0,0,0,0) { 10 13 22 23 27 28 31 34 36 38 39 42 44 48 }					
2. H1V2U272007				22 (1,0,0,0,0) { 12 }					
3. S39621				22 (1,0,0,0,0) { 26 }					

第一列为文件名；MassI 和 MassA 分别为肽段的实验分子质量和理论分子质量；Xcorr 为相关性分析的分数(Score)。分数越高，表明实验数据和理论数据越接近；DelCn 为相关性最高的两个检索结果之间的 Xcorr 差值；Sp 为 MS/MS 丰度最高的 200 个离子与理论离子比较的初步分数；RSp 为由初步分数获得的序列的排列顺序；Ions 为实验所检测到的离子和理论应有的离子；reference 为肽序列的参考号；MH^+ 为实验肽段离子分子质量，Sequence 为肽序列。表中还列出了综合检索到的蛋白质，及其排列顺序和分数。

(三) 序列标签法鉴定蛋白质

Mann 等发展了根据序列标签来搜寻肽段的方法[59]。在 MS/MS 或 PSD 对肽段的序列分析中，通常获得的是部分序列，特别是肽链中段的序列，根据肽段的部分序列和序列两端的肽质量，可以对肽段进行鉴定。可使用的软件有 PepSea。

（四）使用检索工具的注意事项

参数的选择是重要的，会影响到蛋白质鉴定的结果，需要注意的有下述几个方面。

①选择可能遗漏的酶切位点数，一般为 1～2 个，若选择太多，将导致理论库的内容增多，降低了准确性。

②修饰种类的选择，一般半胱氨酸会被烷基化，属于固定修饰。可变修饰也不宜选择太多，因为会提高匹配的难度。

③肽分子质量误差范围，这是影响检索结果的关键因素之一，而且与实验图谱质量相关，是不容易把握的参数，若实验误差较大，而用于检索的误差范围小，就得不到分数较高的检索结果；若实验误差小，用于检索的误差范围大，则匹配的结果太多，同样影响查询结果的可信度。

④用于查库的分子质量，尤其是 MALDI-TOF-MS 图谱中的信号很庞杂，肽信号、杂质、基质等混杂后可能到达上万个峰，而由于检索查库的信号只有几个到几十个，往往会选择分辨率和强度较好的信号，作为用于查库的分子质量。但即便如此，仍然有一些“假峰”，如蛋白酶自酶解峰、角蛋白污染后的酶解峰。因此，需要在查库时进行甄别，消除这些干扰。

（五）蛋白质鉴定的结果评价

通过上述方法在数据库中可以查到蛋白质的全序列，以及质谱检测肽段的覆盖率(coverage)。覆盖率越高表明检索结果越可靠。一般覆盖率为 20%～50%。在有 MS/MS结果的情况下，由于蛋白质序列有很高的专一性，若覆盖率达到 20%～50%，根据分子质量和肽序列所查询结果将是相当可靠的。查询到蛋白质全序列后，通常应根据其精确分子质量和等电点，与蛋白质在双向电泳图谱中的位置进行对比，进一步确证蛋白质的鉴定结果。

胶内蛋白质酶解后，并不是所有的肽段都可以检测到。至少有如下因素会影响到蛋白质的覆盖率：一是蛋白质酶解效率和酶解肽段的获得率。这由蛋白质本身的性质决定，采用酶解与化学裂解(如 CNBr)共同作用的方法，可提高覆盖率[60]；二是蛋白质被修饰的情况。这类蛋白质会出现质量迁移，导致与数据库中的理论肽谱不符合，而未被确认；三是蛋白质很微量或杂质较多，质谱分析的信噪比低，有些分子离子检测不到，因而无法与数据库匹配。

目前质谱法检测的通常是经考马斯亮蓝染色的蛋白质，对分析银染蛋白质仍有一定困难。由于银染蛋白质十分微量，而且银染会导致蛋白质的很多位点被不可逆地修饰和封闭，酶解十分困难。

三、多肽从头测序

仅根据多肽在质谱中的碎片信息解析未知序列，目前还存在困难：一是因为质谱仅获得碎片，无法分辨 b 离子和 y 离子；二是由于除了 b 离子和 y 离子以外，其他离子的存在对序列解析有干扰；三是由多肽断裂不完全，造成离子系列不连续，难以确定某些氨基酸。

目前常用的质谱测序方法大多依赖于蛋白质数据库。测定未知肽序列，称为 *de novo* sequencing，即所谓“从头测序”。从头测序可根据质谱碎片图用软件推导，这一方法将在生物信息学一章中详细阐述。在此介绍化学衍生标记法，化学衍生标记法有如下几种策略。

1. 质量偏移法

将多肽的 N 端或 C 端标记，通过质量偏移区分 b 离子和 y 离子。有人采用酯标记肽 C 端，通过质量偏移区分 b 离子和 y 离子[61]。Cagney 等人[62]则采用 *O*-甲基异脲将胰蛋白酶解肽段中的 Lys 衍生成高精氨酸（homoarginine），与未衍生的肽段 1∶1 混合后，通过 C 端 Lys 的质量偏移，确定 C 端离子。还发展了稳定同位素掺入法，Horiuchi 等[63]在进行蛋白质胰蛋白酶酶解时，用 1∶1 的 H_2O^{18}/H_2O^{16} 作为溶剂，在酶解肽段的 C 端同时有 O^{18} 和 O^{16}，在形成碎片时，同时有这两种氧原子的离子为 C 端离子（图 16.11）。

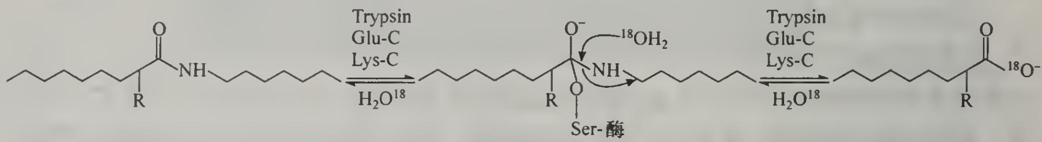
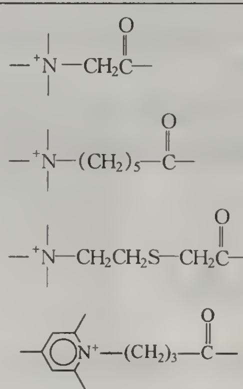


图 16.11　H_2O^{18} 中蛋白质胰蛋白酶酶解过程

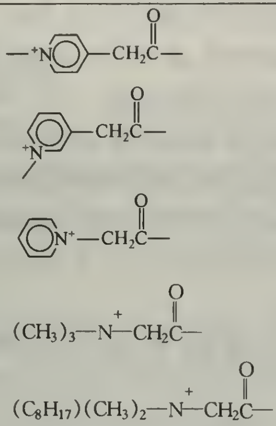
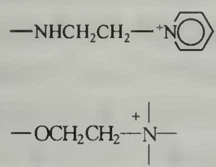
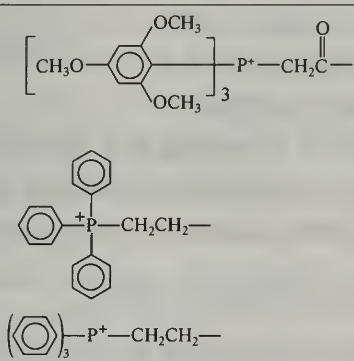
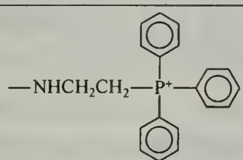
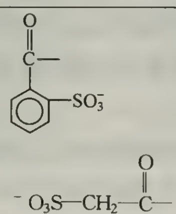
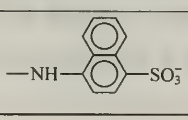
2. 单向离子增抑法

除了质量偏移法以外，另一种思路是在形成质谱图时就只看到单向离子（b 离子或 y 离子），只需从一端便可推导出肽序列。主要的方法是将肽的 N 端或 C 端衍生使其电荷达到增强或削弱，引入正电荷和负电荷均可[64]。在正离子检测模式下，引入带正电荷的基团，可使某一端的离子得到增强，如引入四价季铵或鏻基团；若引入带负电荷的基团，则是某一端的离子减弱甚至完全被抑制，如引入磺酸基团。反之，在负离子检测模式下，引入负电荷则增强，引入正电荷就抑制某一端离子，表 16.7 为目前所采用的衍生基团及其原理。

表 16.7　目前所采用的衍生基团及其原理

衍生基团性质	基团结构	衍生部位	引入电荷性质	文献
quaternary ammonium（季铵）		N 端	正电	[65～70]

续表

衍生基团性质	基团结构	衍生部位	引入电荷性质	文献
				
quaternary ammonium （季铵）		C 端	正电	[71,72]
quaternary phosphonium （鏻基团）		N 端	正电	[73,74]
quaternary phosphonium （鏻基团）		C 端	正电	[74]
Sulfonic acid group （磺酸基）		N 端	负电	[75]
Sulfonic acid group （磺酸基）		C 端	负电	[76]

3. 稳定同位素标记的氨基酸掺入法

由于体外的化学修饰有一定的缺陷(如反应效率问题),普遍有副反应而增加了图谱解析的难度等。另外,一些研究组尝试在细胞培养时,就将已将用同位素标记的氨基酸掺入,因此,在合成出的蛋白质中,就带有同位素标记的氨基酸。Gu[77]等人选择 50% 4 个 D 标记的 Lys(Lys-d4)掺入,使所有蛋白质中的 Lys 都以等量的 Lys-d0 和 Lys-d4 形式存在,由于胰蛋白酶酶切位点即在 Lys 的 C 端。因此,在蛋白质酶切图谱中,形成了 C 端被差异标记的肽段,从而达到了分辨 b 离子和 y 离子的目的。这一方法的最大缺陷在于仅能用于培养细胞,而不能用于组织。

第四节　质谱前蛋白质或多肽样品制备方法和关键步骤

(一) 概　况

质谱前蛋白质或多肽样品准备主要根据纯化系统是凝胶系统还是非凝胶系统,分析系统是 MALDI-MS 还是 ESI-MS,其中又以凝胶系统的 MALDI-MS 或 ESI-MS 样品准备技术要求相对较高,非凝胶系统的 MALDI-MS 或 ESI-MS,无论是完整蛋白质分析还是混合蛋白质(包括单个纯蛋白质)酶解后质谱的样品准备技术要求在本书其他章节将有详细描述。而 HPLC 与 ESI-MS 的联用,样品的脱盐、浓缩也基本实现在线和自动化。目前只有普遍采用的从凝胶到 MALDI-MS 系统的样品准备技术还多少依赖人工操作,在此重点介绍质谱分析前样品准备的整体思路和具体步骤。

凝胶系统一般为 1-D gel 或 2-D gel,胶上分离的蛋白质点可以转移到膜上(如 PVDF 膜)进行膜上酶解,也可以胶内酶解,虽然二者同样有效,但胶内酶解更为方便,且灵敏度较高,有利于低丰度蛋白质的鉴定以及后续的肽段碎片分析(如 PSD 或 MS/MS)[78]。

(二) 质谱前的样品纯化、浓缩

虽然 MALDI-MS 可以容忍一定限度的盐分、去污剂等,但对样品结晶有很大影响以及在离子化过程中形成加合物,将增加数据处理的复杂度,也会降低信噪比,影响结果的可靠性,Zip Tip 是纯化、浓缩蛋白质和肽段的理想工具,为蛋白质组学实验室所必备,关于具体的实验步骤,详见相关章节。

当前关注的焦点是 MALDI-MS 质谱前的样品处理的自动化及怎样提高整个样品前处理的效率。在设备方面已有一些改进,如 Millipore 公司的 ZipPlate Micro-SPE Plate,蛋白质酶解、脱盐、浓缩样品在同一盘中,无需样品转移等繁琐步骤,减少了损失,提高了蛋白质鉴定的成功率和氨基酸的序列覆盖率。

为了满足蛋白质组学中低丰度蛋白质的检测与分析,新的自动化、在线的样品准备方案不断得到发展,瑞典 Lund 大学 Marko-Varga 等人[79,80]做了有益的尝试,他们研制开发的微量纯化硅芯片技术(silicon microextration chip,SMEC)不仅减少了人工操作的步骤,而且减少了层析填料的使用量,并且可重复使用,节约了实验费用,SMEC 甚至可以作为固化微芯片酶反应器(immobilized microchip enzyme reaction,IMER)(与蛋白质纯化中的

柱上复性有异曲同工之妙)，同时对样品洗脱时产生的散逸问题进行了研究。

（三）染色方法与质谱分析的兼容性[81]

1. 传统的染色方法

在蛋白质组学领域中广泛应用的是考马斯亮蓝染色法和银染法。考马斯亮蓝染色法较为灵敏，可以达到微克级检测水平；具有线性的动力学范围，非常适合定量，这一点相当重要；易于使用并且与质谱分析的兼容性相当出色，这些都使其成为常规染色的首选。遗憾的是，微克级的检测水平使它无法检测到相当多的低丰度蛋白质，日益成为蛋白质组学的瓶颈。

银染法比考马斯亮蓝染色法灵敏约 100 倍，检测限度可达纳克级水平，被广泛应用于各蛋白质组学实验室，但是并非完全符合线性动力学(仅在 100～150ng 范围内符合)导致蛋白质差异显示的不准确性，给后续分析及鉴定带来一定的实际困难，实验操作中使用的银离子及戊二醛影响胶内蛋白质酶解，降低了蛋白质质谱分析的氨基酸序列覆盖率。

我们实验室蛋白质(考马斯亮蓝染色)鉴定的成功率为 60%～80%，远高于蛋白质(银染)鉴定的成功率(为 30%～60%)。鉴于此，我们也在不断发展新的方法，以提高银染鉴定的成功率，改进的无戊二醛银染法[82]和 H_2O_2[83]脱色法将会降低不利因素带来的影响。

2. 荧光染色法[84,85]

虽然荧光染色应用到蛋白质组学研究的时间不长，但其突出的优点和广泛的应用前景冲击着传统的染色方法，从事蛋白质组学实验的研究人员将是最大的受益者。荧光染料有多种，其中以 Sypro 染料最为引人注目，荧光染色法与银染法同样具有高灵敏度，但克服了银染法的一些缺点，染料与蛋白质非共价结合，减少了对酶解及质谱分析的影响，极大地提高了氨基酸序列覆盖率，同时具有宽广的线性动力学范围(2～2 000ng)，这使得低丰度蛋白质的检测和精确的蛋白质定量更为准确、简便。但是由于荧光染色法要求使用昂贵的仪器和试剂，延缓了其在蛋白质组学中的推广速度。

（四）关于酶自解峰和角蛋白的污染

首先，要尽量降低自酶解，减少角蛋白的污染，因它们“额外”的肽段会干扰质谱检测得到的数据的分析，主要的预防措施如下所述。

① 使用 TPCK-Trypsin(测序级)。Promega 和 Roche 公司均有较好产品。

② 在加酶使胶块吸胀后，一定要去掉多余酶液，不要认为这样会影响酶解的效率。

③ 酶解时间不宜过长，不要超过 24h。

④ 电泳的玻璃板、染胶的器皿等用前应洗净。

⑤ 所用试剂和溶液的纯度要高，以便达到质谱分析的要求。

⑥ 尽量使用干净的乳胶手套和洁净的 EP 管及 Millipore H_2O。

⑦ 尽量减少不必要的操作步骤，简化操作程序。

⑧ 要有一个洁净的操作环境。

尽管如此,有时自酶解和角蛋白的污染还是难以避免的,那么有必要在数据存储时将其剔除,提高搜库的有效性和准确性,同时它们也可以作为内标准物校准仪器(Porcine Trypsin:842.5100 和 2211.1046,$M+H^+$),相对外标准物有更高的精确度。表 16.8 和表 16.9 分别表示猪和牛胰蛋白酶自酶解理论肽谱(中性,相对分子质量值)。猪和牛胰蛋白酶的全序列参考 http://biochem.roche.com/pack-insert/1418475a.pdf 以及文献[86]。

常见的对质谱影响较大的角蛋白有以下 4 种:角蛋白 1(皮肤)、角蛋白 2E(头皮屑)、角蛋白 9(皮肤)、角蛋白 10(头皮屑)。详情请参见网址 http://www.matrixscience.com/ Help Topic Index:Protein chemistry 中 autolysis 和 contaminants 选项。

表 16.8 猪胰蛋白酶自酶解理论肽谱

相对分子质量	位置	#MC	肽序列
3012.3164	171~200	0	DSCQGDSGGPVVCNGQLQGI VSWGYGCAQK
2282.1729	70~89	0	IITHPNFNGNTLDNDIMLIK
2210.0967	50~69	0	LGEHNIDVLEGNEQFINAAK
2157.0234	150~170	0	SSYPGQITGNMICVGFLEGGK
1767.7920	108~125	0	SCAAAGTECLISGWGNTK
1735.8352	209~223	0	VCNYVNWIQQTIAAN
1468.7232	126~139	0	SSGSSYPSLLQCLK
1044.5564	90~99	0	LSSPATLNSR
1005.4801	140~149	0	APVLSDSSCK
905.4970	201~208	0	NKPGVYTK
841.5021	100~107	0	VATVSLPR
514.3227	46~49	0	IQVR
261.1437	44~45	0	SR

表 16.9 牛胰蛋白酶自酶解理论肽谱

相对分子质量	位置	#MC	肽序列
2551.2410	100~125	0	VASISLPTSCASAGTQCLIS GWGNTK
2272.1521	70~89	0	SIVHPSYNSNTLNNDIMLIK
2192.9870	150~170	0	SAYPGQITSNMFCAGYLEGGK
2162.0491	50~69	0	LGEDNINVVEGNEQFISASK
1494.6079	171~186	0	DSCQGDSGGPVVCSGK
1432.7133	187~200	0	LQGIVSWGSGCAQK
1152.5662	126~136	0	SSGTSYPDVLK
1110.5532	209~217	0	VCNYVSWIK
1019.4957	140~149	0	APILSDSSCK
905.4970	201~208	0	NKPGVYTK
804.4090	92~99	0	SAASLNSR
658.3762	44~49	0	SGIQVR
632.3129	218~223	0	QTIASN
362.1988	137~139	0	CLK
259.1896	90~91	0	LK

（五）MALDI-MS 基质的选择

1. 基质的作用及选择

基质的使用是 MALDI 这一软电离技术的关键，各种基质都有相应的应用范围（表 16.10），不同的生物样品只有在相应的基质辅助下，才能更好地离子化，进入 MALDI 质谱分析。

表 16.10　基质的种类及应用范围

基　质	分子质量/Da	应　用
α-氰基-4-羟基肉桂酸（α-cyano-4-hydroxycinnamic acid, HCCA）	189.04	肽（<5 000Da），脂类，碳水化合物
3,5-二甲氧基-4-羟基肉桂酸（sinapinic acid, SA, or trans-3,5-dimethoxy-4-hydroxycinnamic acid）	224.07	肽和大蛋白质（3 000～150 000Da），糖蛋白，膜蛋白
2,5-二羟基苯甲酸（gentisic acid, or 2,5-dihydroxy benzoic acid, DHB）	154.03	肽，蛋白质，碳水化合物，糖蛋白，糖脂，聚合物，有机分子
3-羟基吡啶甲酸（3-hydroxypicolinic acid, HPA）	139.03	寡核苷酸（>3 500Da）
2,4,6-三羟基苯乙酮（2,4,6-trihydroxyacetophenone, THAP）	168.00	寡核苷酸（<3 500Da）
2,6-二羟基苯乙酮（2,6-dihydroxyacetophenone, DHAP）	152.05	糖蛋白，糖肽，磷酸化肽

2. 基质的结晶

在点样前，须注意：配制饱和基质溶液（对于 HCCA，5～10mg/ml），以吸收足够的能量解吸附；基质与被分析物一定要混匀，以利于单分子解吸。基质的结晶状态可以用显微镜或摄像头（仪器配备）监控，各种基质的结晶如图 16.12～图 16.14（各种基质的结晶状态因被分析物的不同可能有所改变）。



图 16.12　HCCA 的结晶

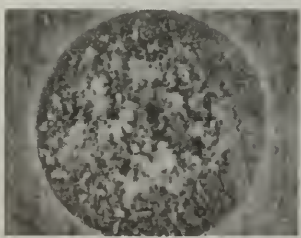


图 16.13　SA 的结晶

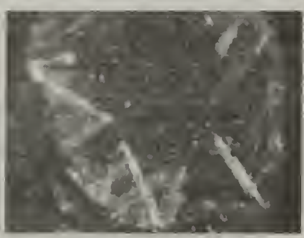


图 16.14　DHB 的结晶

（六）胶内酶解方法[87]

1. 操作步骤

① 用手术刀片切下胶上目标条带（或用自制切点器切下感兴趣的点），置于 EP 管中（胶块切成约 $1mm^3$大小），同时切下空的胶块作对照。

② 加入 200～400μl 100mmol/L NH_4HCO_3/30%ACN 脱色，清洗至透明，去除上清，冻干（脱色液体积过量为好，脱色至透明，不必担心蛋白质的损失，因为完整蛋白质在此条件下很难被洗脱）。

（若为银染）[88]取 30～50μl 30mmol/L $K_3Fe(CN)_6$:100mmol/L $Na_2S_2O_3$ = 1:1（体积比）加入胶块内，洗至蛋白质点棕色消失，去除上清，立刻加 200μl 水终止反应，10min 后，去除上清，再加入 100μl 100mmol/L NH_4HCO_3，静置 20min，去除上清。

③（若是 2-D 胶，则跳至第 8 步）每管加入 90μl 100mmol/L NH_4HCO_3，10μl 100mmol/L DTT，56℃孵育 30min，还原蛋白质（溶液体积应过量，若温度为室温，反应时间相应延长至 1～2h）。

④ 去上清，每管加 100μl 100%ACN，5min 后吸去。

⑤ 每管加入 70μl 100mmol/L NH_4HCO_3，30μl 200mmol/L IAA（现配，避光保存），暗处反应 20min。

⑥ 去上清，每管加入 100μl 100mmol/L NH_4HCO_3，室温反应 15min。

⑦ 去上清，加入 100μl 100%ACN，5min 后吸去，冻干。

⑧ 冻干后，各加入 5μl 2.5～10ng/μl Trypsin 溶液，置于 4℃冰箱 30～60min，使胶块充分吸胀（若仍有剩余液，吸出打掉）。需注意：50ng 酶量基于考染一块胶 1mg 的上样量，12.5ng 基于银染一块胶 100μg 的上样量。上样量增加，酶量同比增加，酶与被分析蛋白质质量比一般为 1:20～1:100[89,90]。

⑨ 再加入 20～30μl 左右 50mmol/L NH_4HCO_3缓冲液（无 Trypsin），pH 7.8～8.0（Promega）。缓冲液体积视胶块体积而定，一般没过胶块 20μl 左右。

⑩ 37℃反应过夜，20h 左右。

⑪ 吸出酶解液，转移至新 EP 管中，原管加入 100μl 60% ACN/0.1%TFA，超声 15min，吸出溶液，并入前次溶液，反复抽提 3 次，合并，冻干。

⑫ 样品制备完成，可以复溶，点样，进行质谱分析。

2. 注意事项

① 若样品种类较多，应将离心管编号，样品对号入管。

② 考染方案各注释均适用于银染方案。

③ 整个操作过程应注意角蛋白等的污染。

④ 佩带一次性乳胶手套，使用洁净的离心管及 Millipore H_2O。

⑤ 不同公司的酶，最适 pH 范围不同，需调节 pH 值。

（七）样品纯化、浓缩 Zip Tip 方案

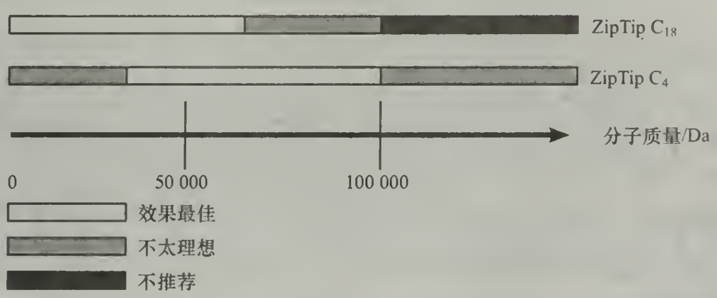
1.Zip Tip 操作步骤

① 润湿 C_{18}介质。
② 平衡 Zip Tip 枪头。
③ 捆绑肽段或蛋白质。
④ 洗去盐分和残留物。
⑤ 洗脱被纯化的肽段或蛋白质。

2. 使用 Zip Tip C_{18}枪头时要求的溶液

润湿液:50% ACN Milli-Q 级水配制。
样品准备液:0.1% TFA Milli-Q 级水配制,样品液 pH<4。
平衡液:0.1% TFA Milli-Q 级水配制。
洗杂液:0.1% TFA Milli-Q 级水配制。
洗脱液:0.1% TFA/50% ACN。

3. 不同介质的 Zip Tip 对不同分子质量范围的肽段或蛋白质的捆绑效率



4. 说明

有关 ZipTip 使用的详细说明,见网址 www. millipore. com/ziptip,一般前面胶内酶解冻干的样品用上面介绍的样品准备液 3μl 复溶(若是多块胶合并产物,适当增加,但不易过多,因为 ZipTip 的容量只有 10μl;若冻干样品不是胶内产物,样品量较多,要适当稀释,因 ZipTip 有一定的载量,C_{18}≥1.0μg,最大为 5.0μg、C_4≥0.5μg,最大为 3.3μg),通过以上介绍的脱盐方案,最终洗脱为 2μl 的样品液进行质谱分析。

因 ZipTip 方案大部分是适用于手工操作的实验,在操作时应注意以下几点:①整个操作过程润湿体积>0.6μl(因填料为 0.6μl);②pH 值<4;③整个过程绝对不能引入气泡;④注意流速,最好速度均匀。

（八）样品-基质的准备方法

1. Dried-Droplet 法

样品-基质的准备方法对 MALDI-MS 的影响非常大,在已经报道的准备方法中,

Dried-Droplet 法被广泛应用(本中心实验室也以此为主)。在这个方法中,饱和的基质溶液与被分析物溶液混合(比例大约为 5 000:1),0.5μl~2μl 混合液被点在靶体上(本中心为 BrukerDaltonics Scout 384),让其自然干燥。

下面是此法的一个具体的例子。

①称取 0.001gCHCA 基质装入洁净的 EP 管中。

②在此 EP 管中加入 200μl 0.1% TFA/50% ACN 溶液,混匀,离心。

③取 1μl 饱和基质溶液与 1μl 样品液(参考浓度约为 1.0pmol/L)混合均匀。

④取 1μl 混合液用加样枪点在靶体盘上(注意不要有气泡,枪头也不要接触靶体表面)。

⑤在洁净的环境中让其自然挥发(在未结晶之前不要晃动)。

2. 其他准备方法

在 Thin-Layer 法[91,92]中,均一的基质层首先在靶体上形成,然后样品加在其上,这种方法有很高的灵敏度、分辨率和精确度。类似地,在 Thick-Layer 方法[93]中,硝化纤维被用作基质辅助物,当均一的硝化纤维-基质层形成后,样品应用其上,这种准备方法抑制了碱性加合物的形成,明显提高了检测的灵敏度,特别是对胶内抽提的蛋白质和肽。Sandwich 法[92]则是另一种"变体",按 Thin-Layer 法基质晶膜形成后,再滴加 TFA、样品和基质。

上述的方案基本都已经流程化,总的来看,MALDI 质谱对样品准备的要求最高,影响因素也最多,基质的选择、基质溶液的组分和浓度、基质溶液的 pH 值以及样品基质共结晶的速度等[94,95]在一定程度上影响 MALDI 质谱分析的结果,要对研究样品达到最好的分析效果,相关操作人员不仅要了解 MALDI 质谱的原理,还要对样品的准备过程有深入的理解并熟练操作。操作人员与样品提供者之间的交流也是相当重要而有益的。以下信息,如分子质量、等电点、样品量(若不是冻干品,溶剂和浓度)和纯度、样品的纯化过程和来源等都是有益的。

(九) 实验室酶解标准化的控制

为了对实验过程进行有效的监控,用 β-casein(Sigma 公司)进行胶内、溶液内酶解与 LCQ(Thermo Finnigan 质谱仪)质谱测试,结果如表 16.11 和表 16.12。

表 16.11 β-casein 胶内酶解 LCQ 质谱测试结果

品名	时间/h	上样量	匹配肽段数(MS/MS)	查库结果
胶内酶解	18	4μg/1.3μl	3(855.1/1119.2/1849.1)	β-casein B
			4(729.9/723.9/762.0/811.9)	β-casein A2 Variant
胶内酶解	18	2μg/1.3μl	4(1964.0/2168.6/2334.4/2462.6)	β-casein precursor
			2(855.1/1119.2)	β-casein B
			6(729.9/723.9/811.9/854.1/995.2/2168.6)	β-casein A2 Variant
胶内酶解	18	500ng/0.5μl	3(723.9/762.0/811.9)	β-casein A2 Variant
胶内酶解	18	250ng/0.5μl	2(762.0/811.9)	β-casein A2 Variant

注:关于 4μg/1.3μl,"/"之前代表 SDS-PAGE 电泳的上样总量,"/"之后代表胶内酶解、ZipTip 纯化后质谱的上样体积。

表 16.12 β-casein 溶液酶解 LCQ 质谱测试结果表

品名	时间/h	上样量/fmol	匹配肽段数(MS/MS)	查库结果
溶液酶解	12	50	3(729.9/723.9/811.9)	β-casein A2 Variant
溶液酶解	16	50	4(729.9/723.9/811.9/2168.6)	β-casein A2 Variant
溶液酶解	21	50	6(611.9/627.8/729.9/723.9/811.9/2168.6)	β-casein A2 Variant
			2(1964.0/2168.6)	β-casein precursor
溶液酶解	12	500	5(729.9/723.9/762.0/811.9/2168.6)	β-casein A2 Variant
溶液酶解	16	500	6(729.9/723.9/762.0/811.9/2043.4/2168.6)	β-casein A2 Variant
			1(2168.6)	β-casein precursor
			2(1119.2/1120.4)	β-casein B
溶液酶解	21	500	7(627.8/729.9/723.9/762.0/811.9/995.2/2168.6)	β-casein A2 Variant
			2(1964.0/2168.6)	β-casein precursor
			1(1119.2)	β-casein B

实验结果表明,β-casein 胶内酶解以 18h、2μg 上样量为好,结果 6 段肽 MS/MS 匹配较好。β-casein 溶液酶解以 21h、500fmol 为好,且有 7 段肽 MS/MS 匹配较好。以上结果经多次实验均鉴定 β-casein (Sigma 公司)以 β-casein A2 Variant 为主要成分。在实验过程中用标准品进行全程对照,监控实验过程,以提高实验结果的可靠性。

(十)小　结

质谱前样品的制备也要根据所用的仪器和研究者的目的而定,是 ESI 质谱还是 MALDI 质谱,是鉴定蛋白质还是对蛋白质的功能进行研究。磷酸化蛋白和糖蛋白的样品制备异于上述常规样品制备方法,本书另有详述。

总的来说,质谱分析对蛋白质的总体要求是“质”和“量”的概念,无论是单个纯的样品还是复杂的混合样品都必须有品质的保证,尽可能避免污染、提高被分析物的纯度,减少各种非蛋白质组分,例如不挥发盐、去污剂等;还有绝对量的满足,尽可能达到各类质谱仪常规检测分析的下限,以提高蛋白质鉴定的成功率和实验结果分析的可靠性。

(史　律　曾　嵘)

参 考 文 献

1 McLafferty FW, Fridriksson EK, Horn DM *et al*. Techview: biochemistry. Biomolecule mass spectrometry. Science, 1999, 284: 1289～1290

2 Costello CE. Bioanalytic applications of mass spectrometry. Curr Opin Biotechnol, 1999, 10: 22～28

3 Aebersold R. Mass spectrometry of proteins and peptides in biotechnology. Curr Opin Biotechnol, 1993, 4: 412～419

4 Cho A, Normile D. Nobel prize in chemistry. Mastering macromolecules. Science, 2002, 298: 527～528

5 Biemann K. Mass spectrometry of peptides and proteins. Annu Rev Biochem, 1992, 61: 977～1010

6 Fenn JB, Mann M, Meng CK *et al*. Electrospray ionization for mass spectrometry of large biomolecules. Science, 1989, 246: 64～71

7 Wilm M, Shevchenko A, Houthaeve T *et al*. Femtomole sequencing of proteins from polyacrylamide gels by nano-electrospray mass spectrometry. Nature, 1996, 379: 466～469

8 Tanaka K, Waki H, Yutaka I *et al*. Protein and polymer analyses up to m/z 100,000 by laser ionization time-of-flight mass spectrometry. Rapid Commun Mass Spectrom, 1988, 2:151~153

9 Hillenkamp F, Karas M, Beavis RC *et al*. Matrix-assisted laser desorption/ionization mass spectrometry of biopolymers. Anal Chem, 1991, 63:1193A~1203A

10 Karas M, Hillenkamp F. Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. Anal Chem. 1988, 60:2299~2301

11 Cramer R, Richter WJ, Stimson E *et al*. Analysis of phospho- and glycopolypeptides with infrared matrix-assisted laser desorption and ionization. Anal Chem, 1998, 70:4939~4944

12 Niu S, Zhang W, Chait BT. Direct comparison of infrared and ultraviolet wavelength matrix-assisted laser desorption/ionization mass spectrometry of proteins. J Am Soc Mass Spectrom, 1998, 9:1~7

13 Wei J, Buriak JM, Siuzdak G. Desorption-ionization mass spectrometry on porous silicon. Nature, 1999, 399:243~246

14 Kruse RA, Rubakhin SS, Romanova EV *et al*. Direct assay of Aplysia tissues and cells with laser desorption/ionization mass spectrometry on porous silicon. J Mass Spectrom, 2001, 36:1317~1322

15 Covey TR, Huang EC, Henion JD. Structural characterization of protein tryptic peptides via liquid chromatography/mass spectrometry and collision-induced dissociation of their doubly charged molecular ions. Anal Chem, 1991, 63:1193~1200

16 Mock K, Hail M, Mylchreest I *et al*. Rapid high-sensitivity peptide mapping by liquid chromatography-mass spectrometry. J Chromatogr, 1993, 646:169~174

17 Davis MT, Stahl DC, Hefta SA *et al*. A microscale electrospray interface for on-line, capillary liquid chromatography/tandem mass spectrometry of complex peptide mixtures. Anal Chem, 1995, 67:4549~4556

18 Davis MT, Lee TD. Rapid protein identification using a microscale electrospray LC/MS system on an ion trap mass spectrometer. J Am Soc Mass Spectrom, 1998, 9:194~201

19 McGinley MD, Davis MT, Robinson JH *et al*. A simplified device for protein identification by microcapillary gradient liquid chromatography-tandem mass spectrometry. Electrophoresis, 2000, 21:1678~1684

20 Licklider LJ, Thoreen CC, Peng J *et al*. Automation of nanoscale microcapillary liquid chromatography-tandem mass spectrometry with a vented column. Anal Chem, 2002, 74:3076~3083

21 Natsume T, Yamauchi Y, Nakayama H *et al*. A direct nanoflow liquid chromatography-tandem mass spectrometry system for interaction proteomics. Anal Chem, 2002, 74:4725~4733

22 Yost RA, Boyd RK. Tandem mass spectrometry: quadrupole and hybrid instruments. Methods Enzymol, 1990, 193:154~200

23 Biemann K. Sequencing of peptides by tandem mass spectrometry and high-energy collision-induced dissociation. Methods Enzymol, 1990, 193:455~479

24 McLuckey SA, Van Berkel GJ, Goeringer DE *et al*. Ion trap mass spectrometry. Using high-pressure ionization. Anal Chem, 1994, 66:737A~743A

25 Marshall AG, Hendrickson CL, Jackson GS. Fourier transform ion cyclotron resonance mass spectrometry: a primer. Mass Spectrom Rev, 1998, 17:1~35

26 Gorshkov MV, Pasa-Tolic L, Bruce JE *et al*. A dual-trap design and its applications in electrospray ionization FTICR mass spectrometry. Anal Chem, 1997, 69:1307~1314

27 Hayes RN, Gross ML. Collision-induced dissociation. Methods Enzymol, 1990, 193:237~263

28 Schreiner M, Strupat K, Lottspeich F *et al*. Ultraviolet matrix assisted laser desorption ionization-mass spectrometry of electroblotted proteins. Electrophoresis, 1996, 17:954~961

29 Sutton CW, Wheeler CH, Sally U *et al*. The analysis of myocardial proteins by infrared and ultraviolet laser desorption mass spectrometry. Electrophoresis, 1997, 18:424~431

30 Eckerskorn C, Strupat K, Schleuder D *et al*. Analysis of proteins by direct-scanning infrared-MALDI mass spectrometry after 2D-PAGE separation and electroblotting. Anal Chem, 1997, 69:2888~2892

31 Loo RR, Cavalcoli JD, VanBogelen RA *et al*. Virtual 2-D gel electrophoresis: visualization and analysis of the E. coli

proteome by mass spectrometry. Anal Chem, 2001, 73(17): 4063～4070

32 Cohen SL, Chait BT. Mass spectrometry of whole proteins eluted from sodium dodecyl sulfate-polyacrylamide gel electrophoresis gels. Anal Biochem, 1997, 247: 257～267

33 Nesatyy VJ, Ross NW. Recovery of intact proteins from silver stained gels. Analyst, 2002, 127: 1180～1187

34 Schuhmacher M, Glocker MO, Wunderlin M *et al*. Direct isolation of proteins from sodium dodecyl sulfate-polyacrylamide gel electrophoresis and analysis by electrospray-ionization mass spectrometry. Electrophoresis, 1996, 17: 848～854

35 Ehring H, Stromberg S, Tjernberg A *et al*. Matrix-assisted laser desorption/ionization mass spectrometry of proteins extracted directly from sodium dodecyl sulphate-polyacrylamide gels. Rapid Commun Mass Spectrom, 1997, 11: 1867～1873

36 Wilkins MR *et al*. In: Proteome Research: New frontiers in Functional Genomics, MR Wilkins *et al*., eds. Springer, Berlin, 1997, 35～91

37 Hunt DF, Yates JR 3rd, Shabanowitz J *et al*. Protein sequencing by tandem mass spectrometry. Proc Natl Acad Sci USA, 1986, 83: 6233～6237

38 Chaurand P, Luetzenkirchen F, Spengler B. Peptide and protein identification by matrix-assisted laser desorption ionization (MALDI) and MALDI-post-source decay time-of-flight mass spectrometry. J Am Soc Mass Spectrom, 1999, 10: 91～103

39 Gevaert K, Demol H, Martens L *et al*. Protein identification based on matrix assisted laser desorption/ionization-post source decay-mass spectrometry. Electrophoresis, 2001, 22: 1645～1651

40 Shevchenko A, Loboda A, Shevchenko A *et al*. MALDI quadrupole time-of-flight mass spectrometry: a powerful tool for proteomic research. Anal Chem, 2000, 72: 2132～2141

41 Krutchinsky AN, Kalkum M, Chait BT. Automatic identification of proteins with a MALDI-quadrupole ion trap mass spectrometer. Anal Chem, 2001, 73: 5066～5077

42 Medzihradszky KF, Campbell JM, Baldwin MA *et al*. The characteristics of peptide collision-induced dissociation using a high-performance MALDI-TOF/TOF tandem mass spectrometer. Anal Chem, 2000, 72: 552～558

43 D'Agostino PA, Hancock JR, Provost LR *et al*. High resolution electrospray mass spectrometry with a magnetic sector instrument: accurate mass measurement and peptide sequencing. Rapid Commun Mass Spectrom, 1995, 9: 597～603

44 Martin SE, Shabanowitz J, Hunt DF *et al*. Subfemtomole MS and MS/MS peptide sequence analysis using nano-HPLC micro-ESI fourier transform ion cyclotron resonance mass spectrometry. Anal Chem, 2000, 72: 4266～4274

45 Bahr U, Karas M, Kellner R. Differentiation of lysine/glutamine in peptide sequence analysis by electrospray ionization sequential mass spectrometry coupled with a quadrupole ion trap. RapidCommun Mass Spectrom, 1998, 12: 1382～1388

46 Bonetto V, Bergman AC, Jornvall H *et al*. C-terminal sequence determination of modified peptides by MALDI MS. J Protein Chem, 1997, 16: 371～374

47 Hulst AG, Kientz CE. Differentiation between the isomeric amino acids leucine and isoleucine using low-energy collision-induced dissociation tandem mass spectroscopy. J Mass Spectrom, 1996, 31: 1188～1190

48 Wee S, O'Hair RA, McFadyen WD. Side-chain radical losses from radical cations allows distinction of leucine and isoleucine residues in the isomeric peptides Gly-XXX-Arg. Rapid Commun Mass Spectrom, 2002, 16: 884～890

49 Vaisar T, Gatlin CL, Rao RD *et al*. Sequence information, distinction and quantitation of C-terminal leucine and isoleucine in ternary complexes of tripeptides with Cu(Ⅱ) and 2,2'-bipyridine. J Mass Spectrom, 2001, 36: 306～316

50 Patterson DH, Tarr GE, Regnier FE *et al*. C-terminal ladder sequencing via matrix-assisted laser desorption mass spectrometry coupled with carboxypeptidase Y time-dependent and concentration-dependent digestions. Anal Chem, 1995, 67: 3971～3918

51 Thiede B, Wittmann-Liebold B, Bienert M *et al*. MALDI-MS for C-terminal sequence determination of peptides and proteins degraded by carboxypeptidase Y and P. FEBS Lett, 1995, 357: 65～69

52 Gu QM, Prestwich GD. Efficient peptide ladder sequencing by MALDI-TOF mass spectrometry using allyl isothiocyanate. J Pept Res, 1997, 49: 484～491

53 Thiede B, Salnikow J, Wittmann-Liebold B. C-terminal ladder sequencing by an approach combining chemical degradation with analysis by matrix-assisted-laser-desorption ionization mass spectrometry. Eur J Biochem, 1997, 244: 750～754
54 Wang R, Chait BT. N-terminal protein ladder sequencing. Methods Mol Biol, 1997, 64: 175～182
55 Zhang W, Chait BT. ProFound: an expert system for protein identification using mass spectrometric peptide mapping information. Anal Chem, 2000, 72: 2482～2489
56 Perkins DN, Pappin DJ, Creasy DM *et al*. Probability-based protein identification by searching sequence databases using mass spectrometry data. Electrophoresis, 1999, 20: 3551～3567
57 Yates JR 3rd. Database searching using mass spectrometry data. Electrophoresis, 1998, 19: 893～900
58 MacCoss MJ, Wu CC, Yates JR 3rd. Probability-based validation of protein identifications using a modified SEQUEST algorithm. Anal Chem, 2002, 74: 5593～5599
59 Mann M, Wilm M. Error-tolerant identification of peptides in sequence databases by peptide sequence tags. Anal Chem, 1994, 66: 4390～4399
60 van Montfort BA, Canas B, Duurkens R *et al*. Improved in-gel approaches to generate peptide maps of integral membrane proteins with matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. J Mass Spectrom, 2002, 37: 322～330
61 Goodlett DR, Keller A, Watts JD *et al*. Differential stable isotope labeling of peptides for quantitation and de novo sequence derivation. Rapid Commun Mass Spectrom, 2001, 15: 1214～1221
62 Cagney G, Emili A. De novo peptide sequencing and quantitative profiling of complex protein mixtures using mass-coded abundance tagging. Nat Biotechnol, 2002, 20: 163～170
63 Shevchenko A, Chernushevich I, Ens W *et al*. Rapid 'de novo' peptide sequencing by a combination of nanoelectrospray, isotopic labeling and a quadrupole/time-of-flight mass spectrometer. Rapid Commun Mass Spectrom, 1997, 11: 1015～1024
64 Roth KD, Huang ZH, Sadagopan N *et al*. Charge derivatization of peptides for analysis by mass spectrometry. Mass Spectrom Rev, 1998, 17: 255～274
65 Stults JT, Lai J, McCune S *et al*. Simplification of high-energy collision spectra of peptides by amino-terminal derivatization. Anal Chem, 1993, 65: 1703～1708
66 Bartlet-Jones M, Jeffery WA, Hansen HF *et al*. Peptide ladder sequencing by mass spectrometry using a novel, volatile degradation reagent. Rapid Commun Mass Spectrom, 1994, 8: 737～742
67 Stults JT, Halualani R, Wetzel R. Amino-terminal derivatization of peptides yields improved CAD spectra. Proc 37th ASMS Conf Mass Spectrom Allied Topics, Miami Beach, FL, May 21～26, 1989
68 Johnson RS, Martin SA, Biemann K. Collision-induced fragmentation of $(M+H)^+$ ions of peptides. Side chain specific sequence ions. Int J Mass Spectrom Ion Processes, 1988, 86: 137～154
69 Zaia J, Biemann, K. Comparison of charged derivatives for high energy collision-induced dissociation tandem mass spectrometry. J. Am. Soc. Mass Spectrom, 1995, 6: 429～436
70 Vath JE, Biemann K. Microderivatization of peptides by placing a fixed positive charge at the N-terminus to modify high energy collision fragmentation. Int J Mass Spectrom Ion Processes, 1990, 100: 287～299
71 Kidwell DA, Ross MM, Colton RJ. Sequencing of peptides by SIMS from the C or N terminus. Springer Ser Chem Phys SIMS 4, 1984, 36: 412～414
72 Bennett BD, Day RA. A novel C-terminus specific peptide derivatization designed to enhance FAB-MS sensitivity: chemical introduction of quaternary amine functionality specifically to C-terminus via azlactone intermediate. Proc 35th ASMS Conf Mass Spectrom Allied Topics, Denver, CO, May 24～29, 1987
73 Huang ZH, Wu J, Roth KD *et al*. A picomole scale method for charge derivatization of peptides for sequence analysis by mass spectrometry. Anal Chem, 1997, 69: 137～144
74 Wagner DS, Salari A, Gage DA *et al*. Derivatization of peptides to enhance ionization efficiency and control fragmentation during analysis by fast atom bombardment tandem mass spectrometry. Biol Mass Spectrom, 1991, 20: 419～425
75 Keough T, Youngquist RS, Lacey MP. A method for high-sensitivity peptide sequencing using postsource decay matrix-

assisted laser desorption ionization mass spectrometry. Proc Natl Acad Sci USA, 1999, 96: 7131～7136

76 Lindh I, Griffiths WJ, Bergman T *et al*. Charge-remote fragmentation of peptides derivatized with 4-aminonaphthalenesulphonic acid. Rapid Commun Mass Spectrom, 1994, 8: 797～803

77 Gu S, Pan S, Bradbury EM *et al*. Use of deuterium-labeled lysine for efficient protein identification and peptide de novo sequencing. Anal Chem, 2002, 74: 5774～5785

78 Lamer S, Jungblut PR. Matrix-assisted laser desorption-ionization mass spectrometry peptide mass fingerprinting for proteome analysis: identification efficiency after on-blot or in-gel digestion with and without desalting procedures. J Chromatogr B, 2001, 752: 311～322

79 Ekstrom S, Malmstrom J, Wallman L *et al*. On-chip microextraction for proteomic sample preparation of in-gel digestion. Proteomics, 2002, 2: 413～421

80 Bergkvist J, Eksrtom S, Wallman L *et al*. Improved chip design for intergrated solid-phase microextration in on-line proteomic sample preparation. Proteomics, 2002, 2: 422～429

81 Lauber WM, Carroll JA, Dufield DR *et al*. Mass spectrometry compatibility of two-dimensional gel protein stains. Electrophoresis, 2001, 22: 906～918

82 Yan JX, Wait R, Berkelman T *et al*. A modified silver staining protocol for visualization of proteins compatible with matrix-assisted laser desorption/ionization and electrospray ionization-mass spectrometry. Electrophoresis, 2000, 21: 3666～3672

83 Sumner LW, Wolf-Sumner B, White SP *et al*. Silver stain removal using H_2O_2 for enhanced peptide mass mapping by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. Rapid Commun Mass Spectrom, 2002, 16: 160～168

84 Malone JP, Radabaugh MR, Leimgruber RM *et al*. Practical aspects of fluorescent staining or proteomic applications. Electrophoresis, 2001, 22: 919～932

85 Patton WF. A thousands points of light: the application of fluorescence detection technologies to two-dimensinal gel electrophoresis and proteomics. Electrophoresis, 2000, 21: 1123～1144

86 Titani K, Ericsson LH, Neurath H *et al*. Amino acid sequence of dogfish trypsin. Biochemistry, 1975, 14(7): 1385～1366

87 Shevchenko A, Wilm M, Vorm O *et al*. Mass spectrometric sequencing of proteins from silver-stained polyacrylamide gels. Anal Chem, 1996, 68: 850～858

88 Gharahdaghi F, Weinberg CR, Meagher DA *et al*. Mass spectrometric identification of proteins from silver-stained polyacrylamide gel: a method for the removal of silver ions to enhance sensitivity. Electrophoresis, 1999, 20: 601～605

89 Protein Translation and analysis: sequencing grade modified trypsin. Promega Life Science Catalog 2002, 261

90 http://biochem.roche.com/pack-insert/1418475a.pdf

91 Vorm O, Mann M. Improved mass accuracy in matrix-assisted laser desorption/ionization time-of-flight mass spectrometry of peptides. J Am Soc Mass Spectrom, 1994, 5: 955～958

92 Vorm O, Roepstorff P, Mann M. Improved resolution and very high sensitivity in MALDI TOF of matrix surfaces made by fast evaporation. Anal Chem, 1994, 66: 3281～3287

93 Kussman M, Nordhoff E, Rahbek-Nielsen H *et al*. Matrix-assisted laser desorption/ionization mass spectrometry sample preparation techniques designed for various peptide and protein analytes. J Am Soc Mass Spectrom, 1997, 32: 593～601

94 Cohen SL, Chait BT. Influnce of matrix solution conditions on the MALDI-MS analysis of peptides and proteins. Anal Chem, 1996, 68: 31～37

95 Bornsen KO. Influnce of salts, buffers, detergents, solvents, and matrices on MALDI-MS protein analysis in complex mixtures. In: Mass spectrometry of proteins and peptides, Chapman JR ed. New Jersey: Humana Press, 2000. 387～404

第十七章　蛋白质组研究中的定量方法

第一节　介　绍

一、概　述

蛋白质组学的研究目的是以大规模的尺度研究细胞蛋白质的功能。依靠 2-DE 和质谱技术,蛋白质组的研究策略和方法在细胞蛋白质的鉴定上已经比较成熟。但是仅仅进行鉴定并不能提供用以阐明蛋白质功能的全部信息,监控蛋白质的表达水平对弄清细胞体内存在的各种生物进程是非常重要的。因此,定量蛋白质组学作为蛋白质组学研究的重要组成部分被提出来。通常情况下,我们是对两种不同状态细胞的蛋白质差异表达进行比较,对蛋白质的定量不必是检测其在细胞内表达的绝对含量,而只需对相对含量的变化进行准确的定量即可。但是,细胞内各种蛋白质的丰度并不单单在转录水平上被调控,还在翻译水平和翻译后修饰的水平被调控,这使得我们不能通过对体内的 mRNA 的丰度进行检测来间接进行准确的蛋白质定量,这些无疑增加了对蛋白质定量的困难。

当前,蛋白质组研究中比较成熟和可信的定量策略和方法主要有两种。其中之一是传统的建立在 2-DE 基础上的染料染色方法,通过比较蛋白质在不同胶上的染色强度来进行相对定量。对蛋白质进行染色已经不仅仅是为了显示蛋白质的存在,而且还提供了其表达水平的信息。用这类方法准确定量的关键是找到灵敏度高、检测动态线性范围大的染料。传统的考马斯蓝染色和银染已经不能满足准确定量的需求,现在发展了不少新的荧光染料染色技术[1]。例如,利用丙基 Cy3 和甲基 Cy5 分别对两个来自不同蛋白质样品进行荧光标记,以此可在同一块凝胶上显示两个样品的差异。随着最新开发出来的 SYPRO Ruby 等发光金属螯合荧光染料的应用,染色进行差异定量表达的准确性将越来越高。但是这类方法存在着一些弊端。价格不菲的染料和显色设备的投入限制了进一步的应用;目前染料染色的准确性也不尽如人意,增加了后续的操作步骤,而且容易给后续的质谱检测带来干扰。最关键的一点是这种方法立足的 2-DE 平台带来的弊端。2-DE 不能检测出具有极端等电点的、分子质量太大和太小的以及低丰度的蛋白质以及膜蛋白,而且分离的效果不是很绝对,许多单一的蛋白质点包含了一个以上的蛋白质,因此来定量就勉为其难了。

在蛋白质组研究中,第二种定量策略是利用质谱检测技术。质谱仪器并不能像荧光检测仪那样,检测到的信号强度对来自单一样品的蛋白质并不具有定量功能,因而不能对来自相同肽段进行两次质谱检测得到的信号强度进行比较定量。但是如果我们对来自不同样品的肽段标上一个内部标准(internal standards),使得可以识别不同样品来源的肽段,同时标记好的肽段的化学性质基本上是一样的,这样就可以确保同一次质谱扫描中两者的离子化效果是相同的,此时肽质谱峰信号成对出现,质谱峰的信号强度就可以作为定

量的依据。经常用到的这个内部标准就是稳定同位素(如 D、^{13}C 和 ^{15}N 等)标记的小分子,这些内标分子必须满足一定的条件,例如,不同样品来源的肽段标记后化学性质是相同的;如果与高效液相色谱联用,标记后的肽段的保留时间尽可能相同,等等。

二、化学标记定量方法

以质谱检测技术为基础的化学标记定量方法有多种,归纳起来主要有两类,一种是体内标记,另一种是体外标记。

(一)体内标记

细胞在含有稳定同位素的介质中生长,稳定同位素被结合进蛋白质,从而充当了蛋白质定量的内部标准。

1. ^{15}N 体内代谢标记

用富含 ^{15}N 的介质来培养酵母、细菌和哺乳类等细胞,^{15}N 被掺入合成的蛋白质中,从而充当其定量的内部标准[2,3],这是一种普遍标记的方法。Oda 等人把这方法用于了对酶表达和磷酸化修饰相对变化的定量[2]。Oda 将野生型酵母细胞培养物在富含轻氮同位素的正常介质(^{14}N,99.6%,^{15}N,0.4%)中生长,突变型酵母细胞培养物放在富含重氮同位素的相同介质($^{15}N>96\%$)中生长。经过一个适当的生长期后合并细胞培养物,提取蛋白质,通过高效液相色谱和 SDS-PAGE 进行分离,胶内酶切感兴趣的蛋白质点,然后进行质谱鉴定和定量。每结合一个 ^{15}N,该肽将比正常 ^{14}N结合的相同肽大一个质量单位,使得不同细胞状态来源的肽段得到区分,肽质谱峰信号成对出现,从而根据质谱峰的信号强度进行准确的定量。但是这种体内标记策略有一个明显的缺点,就是由于相同肽段的不同同位素标记形式之间的质量变化依赖于氮原子数目,因此,对未知的蛋白质难以进行定量。定量的蛋白质的序列事先必须已知,这样才能知道稳定同位素的结合增加了多少质量单位,才能在质谱图上找到这样的一对峰,所以,不能在鉴定蛋白质之前进行定量。而在实际操作中,我们一般首先想要知道的是定量结果,然后才是对表达量有差异的蛋白质进行鉴定。因此,这种体内稳定同位素普遍标记的方法并不能广泛应用。

2. 稳定同位素标记的必需氨基酸体内标记(SILAC 法)

体内标记定量蛋白质表达还有一种策略就是在培养介质中加入稳定同位素标记的必需氨基酸,这主要用于高等动物细胞中蛋白质的定量以及鉴定。高等动物细胞的生长中需要一些必需氨基酸的摄入,如赖氨酸(Lys)、亮氨酸(Leu)、苯丙氨酸(Phe)等,这些氨基酸细胞自身无法合成,需要从外界摄入补充。因此,在培养细胞时可以在培养介质中特异地加入用稳定同位素标记的某一种必需氨基酸,在经过适当的培养时间后,细胞中合成的含有这类氨基酸的蛋白质几乎都掺入了标记的氨基酸,收获不同同位素标记的不同生长状态下的细胞,混合裂解,再做质谱,就可以从质谱图上得到蛋白质差异表达量的信息。Ong 等首先把这种蛋白质定量策略称为 SILAC(stable isotope labeling by amino acids in cell culture,细胞培养物中含有稳定同位素的氨基酸标记),并且利用这种策略通过用 D

标记的 Leu 研究了肌肉细胞分化过程中蛋白质的差异表达[4]。Gu 等也用 D 标记的 Lys 研究了人皮肤成纤维细胞在热休克情况下蛋白质表达的差异[5]。相比较而言，用 Lys 来标记明显优于其他的必需氨基酸，因为当用胰蛋白酶(trypsin)酶切时，保证了含有 Lys 残基的每个肽段都只含有一个标记物，从而避免了未知肽段用^{15}N标记所带来的质量变化不清的类似问题，而且有利于蛋白质做 MS/MS 鉴定时质谱图的解析。

总的来说，体内标记相对于体外标记方法来说具有一定优势，它省去了标记化学反应和分离纯化等步骤，因而减少了样品的损失，有利于低丰度蛋白质的高灵敏度检测。同时，体内标记也存在一些缺点。首先，它不能对组织来源的蛋白质进行分析。其次，培养细胞在富含稳定同位素的介质中生长，这些介质可能会细胞的繁殖和其他生物进程，由此改变蛋白质的表达水平。第三，这种方法受到同位素材料成本的限制，不可能对整个动物体进行标记。最后，对于未知序列肽段，大多数的体内标记都无法确定标记物与肽段之间量的关系，从而使得我们难以确认同一肽段在质谱图上成对出现的不同同位素标记形式。

（二）体外标记

近年来，定量蛋白质组学的发展主要集中在以质谱检测为核心的体外标记上，涌现了一些新的蛋白质定量方法。这些方法消除了体内标记定量带来的缺憾，使样品来源不再局限于培养的细胞，从而具有良好的发展前景，虽然由此增添了化学反应和分离的步骤导致样品的损失。

根据标记的部位不同，我们又可以将体外标记分为以下几种类型。

1. 氨基($—NH_2$)标记

肽段的 N 端和 Lys 残基侧链上是$—NH_2$基团，容易跟含有酰氧基的小分子如琥珀酸酐、*N*-乙酰氧琥珀酰亚胺等发生酰基化反应，从而对肽段进行标记。例如，Chakraborty 等人用正常的 *N*-乙酰氧琥珀酰亚胺和稳定同位素标记的 *N*-乙酰氧-[D3]琥珀酰亚胺研究了 *E. coli* 细胞在 α-半乳糖苷酶激活前后蛋白质表达量的定量变化[6]。

含有稳定同位素的分子特异标记 N 端不仅有利于蛋白质的定量，而且还充当了后续的质谱 MS/MS 扫描中 b 系列离子的标志，从而方便了蛋白质的鉴定。Munchbach 等人于 2000 年首次提出了 N 端稳定同位素特异标记的方法[7](图 17.1)。他们先用 2-DE 分离得到不同条件下生长的 *E. coli* 细胞提取物的蛋白质，选出感兴趣的蛋白质点后切下，Lys 残基在弱碱性条件下被琥珀酸酐酰化，随后蛋白质用 Asp/Glu 蛋白酶酶切。从状态 1 得到的蛋白质 N 端用 H4-Nic-NHS[1-(H4-烟酰氧基)琥珀酰亚胺]特异标记，来自状态 2 的蛋白质则用氘代试剂 D4-Nic-NHS 标记。MALDI-TOF-MS 分析单个肽段的 D4/H4 比率进行蛋白质的相对定量。这套方法的秉承了体内标记的优点，且能够用于所有类型的蛋白质提取物，费用低廉。虽然 Munchbach 在应用这套方法进行蛋白质定量和鉴定时还是建立在 2-DE 分离蛋白质的基础上，实际上它可以对 1-D 胶条和 2-DE 未能完全分离的含多个蛋白质点进行蛋白质定量，从而使我们有条件去研究在 1-DE 上简单分离而不能在 2-DE 上运行的膜蛋白以及低丰度蛋白质。同时对 N 端的特异标记也简化了 MS/MS 质谱图谱的解释。

与特异标记 N 端的$—NH_2$基团相对的是反应试剂特异标记 Lys 残基上的 $ε—NH_2$基

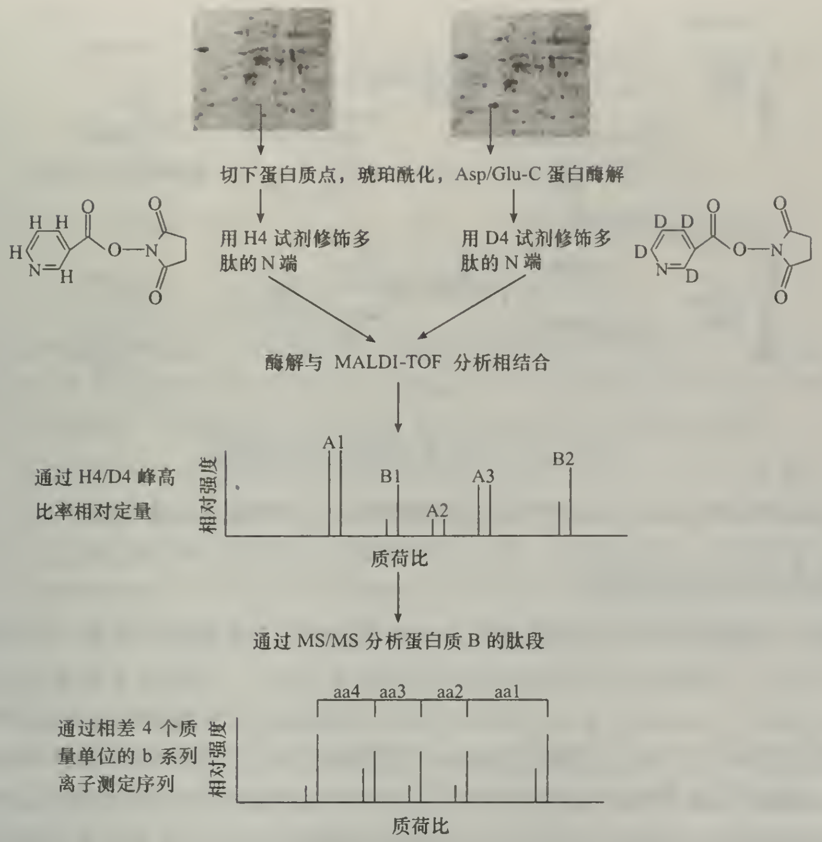


图 17.1　N 端稳定同位素特异标记的方法

团。*O*-甲基异脲(*O*-methylisourea)可以选择性地胍基化 Lys 残基上的 ε—NH_2。Cagney 等首先采用了这种标记方式[8]。但是与以往稳定同位素标记不同的是,他们采用的是一种称之为"质量标记的丰度标签"(mass-coded abundance tagging,MCAT)的标记策略(如图 17.2 所示)。他们的报道打破了以前认为不能只对一份样品进行化学修饰标记来定量的成见,认为这样来做相对定量是可能和准确的。但是由于只对其中一份样品进行标记,来自不同样品的相同肽段在质谱中的离子化效率不一定相同,因此,还需要进一步摸索才能够判断这样的标记定量方法是否可靠和准确。

N 端标记或者说对—NH_2的标记像体内 ^{15}N 代谢标记一样,也属于一种普遍标记的模式。由于通量大的原因,这种模式经常要借助于其他的蛋白质分离和显色技术来对感兴趣的蛋白质进行分析。如采用这种普遍标记的方法来大规模的鉴定和定量蛋白质,则劳动强度非常大。例如,酵母和 *E. coli* 是进行蛋白质组研究的模式生物,按照一个基因对应一个蛋白质的理论,酵母细胞有 6 200 多个蛋白质,经过 trypsin 酶切后将产生 132 768个肽段,平均每个蛋白质有 53 个肽段[9]。但理论上说,每个蛋白质只要有一条肽段就可以对其进行鉴定和定量。因此,要大规模对细胞的整个蛋白质表达量的差异做一个全面的分析,势必要在尽量不放过每个蛋白质的基础上减少质谱分析的工作量。Wang 等人通过联用针对特殊氨基酸组氨酸(His)和半胱氨酸(Cys)残基的亲和色谱分离技术

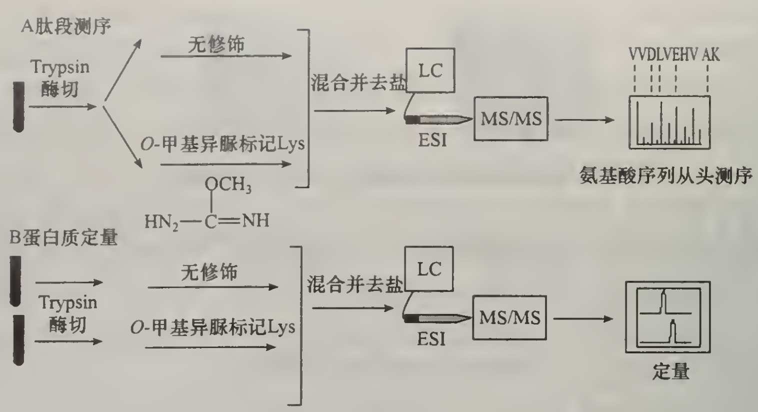


图 17.2　MCAT 蛋白质鉴定和定量策略

和—NH_2标记技术分析了 *E. coli* 中表达的蛋白质,大大降低了分析强度[9]。

2. 羧基(—COOH)标记

肽段的 C 端和所含的天冬氨酸(Asp)和谷氨酸(Glu)残基侧链上的—COOH 羧基也可以进行标记。这种标记方法主要通过两条路线来完成。一条路线主要是通过对这些羧基酯化进行标记。Goodlett 等人已报道用 H 和 D 标记的甲醇酯化标记来定量研究蛋白质表达量的差异[10]。不过这条路线有两个主要缺点:其一是特异性不是很好,在 C 端和 Asp 和 Glu 残基上都有标记,且效率不均。其二是采用的标记条件容易引起天冬酰胺(Asn)和谷氨酰胺(Gln)的去酰胺。另一条路线就是在 H_2O^{18}的溶液中进行蛋白酶切,从而特异性地只在 C 端引入稳定的同位素标记[11]。这种标记像特异标记 N 端一样,其优点是方便了 MS/MS 扫描测序时对 y 系列离子的识别。

3. 巯基(—SH)标记——ICAT 策略

在蛋白质含有的氨基酸残基中,没有哪一个能比半胱氨酸(Cys)残基上的—SH 基团更具有标记反应的特异性。烷基化试剂(如碘乙酸等)能专一地对 Cys 的巯基进行烷基化。如果将稳定同位素引入这些烷基化试剂,那么就可以通过比较质谱图谱上成对出现的质谱峰信号的强度进行相对定量了。由于 Cys 是一种特殊的氨基酸,在蛋白质酶切产生的多肽片段不一定含有 Cys 残基,因此,这种标记方法并不是像以上所述的各种方法那样属于一种普遍标记的模式。如果能将这些烷基化标记的肽段分离出来进行鉴定和定量,无疑将大大减少工作强度。据统计,*E. coli* 和酵母细胞中含有 Cys 残基的蛋白质分别占到蛋白质总量的 85.7% 和 91.6%,而在 trypsin 酶切产生的肽段中,含有 Cys 残基的多肽分离仅占 9.4% 和 9.3%[9],因此,这种标记定量方法与上述标记方法相比,既能监察到足够多的蛋白质,又大大减少了工作量,更适宜大规模操作。

基于这种定量战略,Gygi 等于 1999 年提出了称为“同位素标记的亲和标签”(isotope-coded affinity tag,ICAT)定量策略[12]。

(1) *ICAT 原理和应用步骤*

ICAT 方法建立在一种新型的化学反应试剂——ICAT 试剂的基础上,ICAT 试剂结

构如图 17.3 所示。

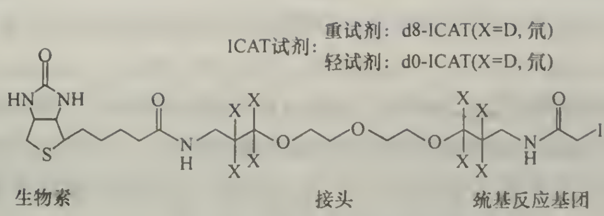


图 17.3　ICAT 试剂的结构

ICAT 试剂包括三部分。①反应基团,可特异地与蛋白质侧链上的 Cys 残基的巯基(—SH)进行反应,使试剂共价结合在蛋白质侧链上。②同位素标记的接头,含有稳定同位素(H 和 D)的标记。③亲和标记的生物素,用于分离标记的蛋白质或多肽。

当 ICAT 试剂中 8 个 X 位置被 H 取代时,称为轻试剂(D_0),当 8 个 X 位置被 D 取代时,称为重试剂(D_8)。因此重试剂比轻试剂的相对分子质量要大 8Da,而且两种试剂的化学性质相同,当分别用两种试剂去标记不同状态下的同一种蛋白质时,不同标记的蛋白质之间的分子质量就相差 8Da,体现在质谱图上就是这对标记蛋白质的肽段离子蜂质荷比(m/z)相差为 8 或 4(当为双电荷肽段离子时),从而能将来自不同样品的同一种蛋白质分离开。

ICAT 标记方法的步骤如下,参见图 17.4。

①将来自一种状态下的细胞的蛋白质混合物还原后用 D_0试剂标记蛋白质侧链上的 Cys 残基;将来自另一种状态下的同种细胞的蛋白质混合物还原后用 D_8试剂标记。

②两种样品混合在一起,用 trypsin 酶切,产生多肽片段,包括标记的和没标记的。

③样品经 SCX-HPLC 两维色谱分离,除去消化酶、去垢剂、还原剂和 ICAT 试剂。

④混合物过抗生物素亲和色谱柱分离,得到只含有同位素标记的多肽混合物。

⑤将分离得到的标记多肽混合物再用 RP-μlC-MS 分离和鉴定。

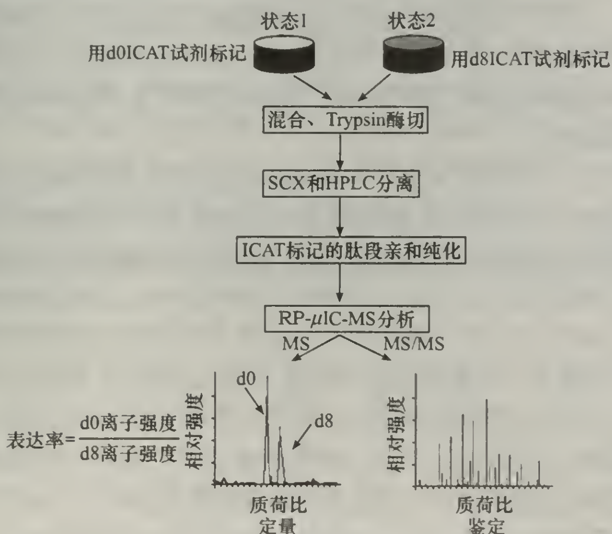


图 17.4　ICAT 方法操作步骤

(2) ICAT 策略的优点和应用

应用 ICAT 策略对蛋白质进行鉴定和定量有着许多的优点，除了能准确进行定量外，还具有以下特征。

①若一种蛋白质酶切产生至少有一个 Cys 残基的肽段的话，就可以对其鉴别和定量。含有 Cys 残基的肽段越多，结果对照得出的结论越准确。

②鉴定的肽段中肯定都含有至少一个 Cys，因此，在进行数据库搜索时就增加了一个有效的约束条件。

③肽段根据标记情况有选择性地得到了富集，从而减少了下游质谱分析的复杂程度。例如，Gygi 等用 ICAT 处理酵母蛋白混合物时，从 344 855 个肽段中富集了 30 619 个标记肽段，大大减少了后期的工作量[12]。

④分离出来的标记多肽可直接与后面的 RP-μlC-MS 分析系统在线耦联操作。

基于此，ICAT 策略现已在蛋白质组研究领域得到了广泛应用。Gygi 等用 ICAT 分析方法首先比较了酵母细胞在不同碳源生长条件下蛋白质表达水平的差异[12]，进而又分析了碳源的改变对细胞内 mRNA 和蛋白质表达水平的影响[13]。Han[14]等用 ICAT 策略分析了用 PMA 处理细胞后微粒体蛋白质的表达变化。Ideker[15]等综合了基因组和蛋白质组分析的概念，描述了一套运用 cDNA 微阵列和 ICAT 策略这些整体研究手段来建立、检验、改进一种细胞生理状态模型的策略，深化了人们对酵母细胞改变碳源后基因表达的认识。

(3) ICAT 策略的缺陷及改进

ICAT 策略并非是十全十美的，它本身也有一些缺点，仍处在不断的发展完善当中。酵母细胞内有 10% 左右的蛋白质上没有 Cys 残基，得不到鉴定[6,9]。ICAT 标记的分子质量是 400Da 左右，对于较小肽段的修饰可能会影响数据库搜索。Zhou 等采用了固相氨基酸同位素标记试剂来代替 ICAT 试剂，降低了标记的分子质量，提高了灵敏度[16]。以 ICAT方法为代表的稳定同位素标记定量策略的缺点是耗时太长和数据存储消耗在量没有改变的蛋白质上，这些蛋白质可能与研究的问题没有太大的关联，新的解决方案是把鉴定和定量分开，串联两套 RP-μlC-MS 分析系统，第一次分析时进行相对定量分析，只有量有变化的肽段才被有选择地送到下游，进行二次鉴定；最近又采用新的质谱系统 MALDI-QTOF 来解决这个问题[17]。

Goshe 等改进了 ICAT 试剂，发展了一种“磷酸多肽同位素亲和标记标签”(PhIAT)技术[18]。PhIAT 试剂可以标记蛋白质上磷酸化的丝氨酸(Ser)和苏氨酸(Thr)残基，用来进行蛋白质磷酸化程度的定量检测，从而为定量研究蛋白质的翻译后修饰提供了新的思路和途径。

第二节　材　料

一、试剂和溶剂

胰蛋白酶(typsin)为 Roche 公司的序列级产品(Sequencing Grade)，乙腈为 Fisher 公

司色谱级产品(Fisher Scientific),HCCA 为 Sigma 公司产品,其余试剂均为 Sigma 公司产品。ICAT 试剂盒为应用生物系统公司的产品(Applied Biosystems)。

细胞裂解液(6mol/L 尿素,1% SDS,100mmol/L DTT,50mmol/L Tris,5mmol/L EDTA,1mmol/L PMSF),溶液 A(6mol/L 尿素,0.05% SDS,50mmol/L Tris,5mmol/L EDTA),蛋白质沉淀液[丙酮:乙醇:乙酸=50:50:0.1(体积比)]。

二、多肽、蛋白质和细胞

标准肽层粘连蛋白(laminin)为 ICAT 试剂盒中自带,标准蛋白质 α-lactoglobulin 为 Sigma 公司产品。其余试剂为国产分析纯。钩端螺旋体菌由上海第二医科大学郭晓奎教授提供。

三、仪　　器

电喷雾质谱(ESI-MS)为 Thermo Finnigan 公司的 LCQ 离子阱质谱系统,基质辅助激光解析离子化-飞行时间质谱(MALDI-TOF-MS)为 Bruker 公司的 REFLEX Ⅲ系统。

第三节　方　　法

1. 标准肽段 Laminin 的 ICAT 标记

将 100μg 标准肽 Laminin 溶于 100μl 溶液 A 中,均等地分为两份,分别加 2μl Reducing Buffer,沸水浴 10min,自然冷却。分别将两管中的溶液移入各一个单位的 D_0 和 D_8 试剂中,避光,室温反应 1h。各取 1μl 混合,进行 MALDI-TOF-MS 分析。

2. 标准蛋白质的同位素标记、酶解和特异性肽段的纯化

称取 1mg 标准蛋白质,溶于 100μl 纯水中。吸取 10μl 分装于 2 个 Eppendorf 管内(control and test),各加 90μl 变性缓冲液(denaturing buffer)和 2μl 还原缓冲液(reducing buffer)于这两管内,置于沸水中水浴 10min,自然冷却,低速离心。分别将两管中的溶液移入各一个单位的 D_0 和 D_8 试剂中,避光,室温反应 1h,低速离心。另取两个 Eppendorf 管,在其中一个加入 50μl D_0 标记样品和 50μl D_8 标记样品;在另一个管内,加入 25μl D_0 标记样品和 50μl D_8 标记样品。(使之质量比分别为 1:1 和 1:2)用 200μl 纯水溶解一份胰蛋白酶(25μg),在两种质量比的样品中各加一半的胰蛋白酶,37℃ 反应,12～16h。组装阳离子交换柱柱架,完毕后,注射 2ml 阳离子交换上样缓冲液。将酶解后的溶液移入两个 5ml的 Eppendorf 管内,加入 2ml 阳离子交换上样缓冲液稀释。以每秒一滴的速度将稀释后的样品注射到阳离子交换柱中,收集流穿液。注射 1ml 阳离子交换上样缓冲液,清洗阳离子交换柱。以每秒一滴的速度注射 500μl 阳离子交换洗脱缓冲液,收集洗脱的肽段。注射 1ml 阳离子交换清洗缓冲液,洗去胰蛋白酶。(另一份不同质量比的样品重复上述离子交换柱层析过程)注射 2ml 阳离子交换上样缓冲液,4℃ 保存过夜,注射 2ml 阳离子交

换贮存液，4℃保存1d以上方可再使用。

组装 Avidin 亲和柱，注射1ml 亲和洗脱缓冲液，再注射2ml 亲和上样缓冲液。在阳离子交换柱的洗脱肽段样品中加入500μl 亲和上样缓冲液。准备3个5ml的 Eppendorf 管，分别编号为a管、b管、c管。以每秒一滴的速度注射中和后的样品，在a管中收集流穿物，注射500μl 亲和上样缓冲液，继续在a管中收集流穿物。注射1ml 亲和清洗缓冲液1，注射1ml 亲和清洗缓冲液2，在b管中收集前500μl。以每秒一滴的速度注射800μl 亲和洗脱缓冲液，弃去前50μl，在c管中收集余下的溶液。注射1ml 亲和洗脱缓冲液。（另一份不同质量比的样品重复上述亲和柱层析过程）注射2ml 亲和上样缓冲液，中和亲和柱。注射2ml 亲和上样缓冲液，4℃保存过夜后，注射2ml 亲和贮存缓冲液，保存1d以上方可再使用。实验中c管中的溶液用于下一步的质谱分析。

3. 细胞抽提总蛋白质的 ICAT 标记和酶解

取钩端螺旋体菌体加入细胞裂解液超声裂解，4℃、15 000r/min 离心1h。取上清，蛋白质定量，采用 Bradford 法。

按5∶1（蛋白质沉淀液∶细胞裂解液＝5∶1，体积比）的比例向定量后的细胞裂解液中加入预冷的沉淀液，置于－20℃。半小时后取出，4℃、15 000r/min 离心1h。去上清，加入70％乙醇（体积与第一步加入的沉淀液相等），洗涤蛋白质多次。4℃、15 000r/min 离心30min，去除上清，将沉淀冻干。

使用适当体积的溶液A溶解冻干的蛋白质，浓度至约10g/L，按所得蛋白质浓度，蛋白质分装成100μg一份，－80℃保存。

各取对照组和实验组100μg 蛋白质，加入 Denaturing Buffer 至总体积为100μl。各加入2μl Reducing Reagent。沸水浴10min，自然冷却。低速离心。将对照组样品移入一个单位的D_0试剂中，混匀。将实验组样品移入一个单位的D_8试剂中，混匀。避光，室温反应2小时。各加1μl 1mol/L DTT，终止反应，室温15min。低速离心。另外在1个 Eppendorf 管内加入95μl 标记后的对照组蛋白质和实验组蛋白质，混匀。取一份 Trypsin（25μg），加200μl Millipore 水溶解加入蛋白质混合物中，37℃（12～16h）。酶解后产物进行质谱分析。

4. MALDI-TOF-MS 分析

使用布鲁克公司的 REFLEX Ⅲ MALDI-TOF 质谱进行分析。用50％乙腈，0.1％TFA 溶解 HCCA，配成饱和溶液溶液，取1μl 与1μl 待分析的 ICAT 标记肽段混合，混匀后取其中1μl 点在加样板上，自然挥干，然后进行 MS 分析。

5. ESI-MS 分析

一维色谱质谱联用在 Finnigan 公司的 LCQ DECA XP 系统上进行。其中反相 HPLC 使用C_{18}柱（150mm length，0.18mm I.D.）其洗脱梯度为：A液为纯水，含有0.1％甲酸；B液为乙腈，含有0.1％甲酸。0～20min，B液2％，20～30min，B液线性升至30％，30～80min，B液线性升至98％，80～81min，B液降至2％。泵流速为150μl/min，分流后为2～3μl/min。质谱条件为：喷雾电压3.5kV，毛细管温度150℃。

MS/MS能量为35%。

二维色谱-质谱联用在Finnigan公司ProteomeX系统。第一相采用阳离子交换柱(150mm length,0.32mm I.D.),第二相采用反相C_{18}柱(150mm length,0.18mm I.D.)。第一相洗脱液为NH_4Cl,浓度梯度为:0,25mmol/L,50mmol/L,75mmol/L,100mmol/L,150mmol/L,200mmol/L,400mmol/L,800mmol/L。反相色谱A液为纯水,含有0.1%甲酸;B液为乙腈,含有0.1%甲酸。洗脱梯度为:0~20min,B液2%,20~80min,B液线性升至98%,80~81min,B液降至2%。质谱条件为:喷雾电压3.5kV,毛细管温度150℃。MS/MS能量为35%。

第四节 结果和讨论

1. ICAT标记标准肽段的MALDI-TOF-MS结果

先使用合成肽Laminin测试方法的可靠性,肽序列为CDPGYIGSR,理论分子质量为967.07Da,含有1个半胱氨酸,标记上轻链和重链ICAT试剂后分子质量应分别增加442.2Da和450.2Da。D_0和D_8分别标记Laminin后,将产物1:1混合,用MALDI-TOF-MS分析。图谱见图17.5。结果显示,重链标记和轻链标记的肽在质谱中的信号强度接近1:1。说明标记和质谱检测定量效果都十分理想。

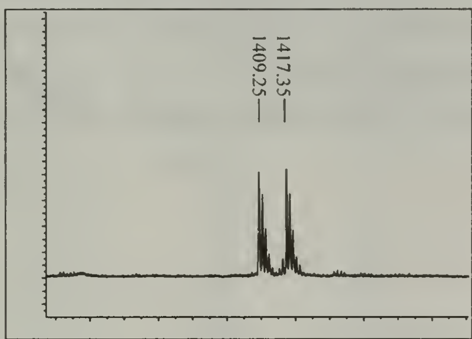


图17.5 D_0和D_8试剂分别标记Laminin的质谱图

2. ICAT标记标准蛋白质的ESI-MS结果

由于简单的合成肽无复杂的空间结构,标记效率高。标记后也无未标记肽干扰,无须纯化。但对于蛋白质,情况会复杂一些。因此,我们进一步选择了标准蛋白质β-lactoglobin进行分析,该蛋白质的理论氨基酸序列是:

MKCLLLALALTCGAQALIVTQTMKGLDIQKVAGTWYSLAMAASDISLLDAQSA PLRVYVEELKPTPEGDLEILLQKWENGECAQKKIIAEKTKIPAVFKIDALNENKVLV LDTDYKKYLLFCMENSAEPEQSLACQCLVRTPEVDDEALEKFDKALKALPMHIRLS FNPTQLEEQCHI(自NCBI,GI:1070647)。经过Trypsin的酶解,理论肽段如表17.1所示。将两份样品分别用D_0试剂和D_8试剂标记后,按1:2比例混合,进行电喷雾质谱分析。

表 17.1　β-lactoglobulin 的理论酶解肽段

分子质量/Da	氨基酸位置	肽序列
2707.3759	31～56	VAGTWYSLAMAASDISLLDAQSAPLR
2647.2023	118～140	YLLFCMENSAEPEQSLACQC LVR
2313.2587	57～76	VYVEELKPTPEGDLEILLQK
2275.2586	3～24	CLLLALALTCGAQALIVTQT MK
1658.7843	165～178	LSFNPTQLEEQCHI
1245.5845	141～151	TPEVDDEALEK
1065.5826	108～116	VLVLDTDYK
1064.4465	77～85	WENGECAQK
916.4734	100～107	IDALNENK
837.4763	158～164	ALPMHIR
674.4235	94～99	IPAVFK
673.3879	25～30	GLDIQK
573.3606	87～91	IIAEK

质谱总离子流图见图 17.6，在对 LCQ 图谱进行分析后，发现保留时间约为 46min 时出现一对信号峰，质荷比分别为1 051.3和1 055.3(图 17.7)。图 17.8 显示的是1 051.3峰的 Zoomscan 图谱，在图中明显可以看到 D_0和 D_8标记肽段的质荷比最小相差是 0.5，就是说两个肽段都是双电荷，肽的实验分子质量应该是 2 100.6Da 和 2 108.6Da。与165～178 位的肽段 LSFNPTQLEEQCHI(理论分子质量 1 658.78Da)被 ICAT 试剂标记后的分子质量(理论分子质量 2 100.98Da 和 2 108.98Da)相符。精度令人满意。

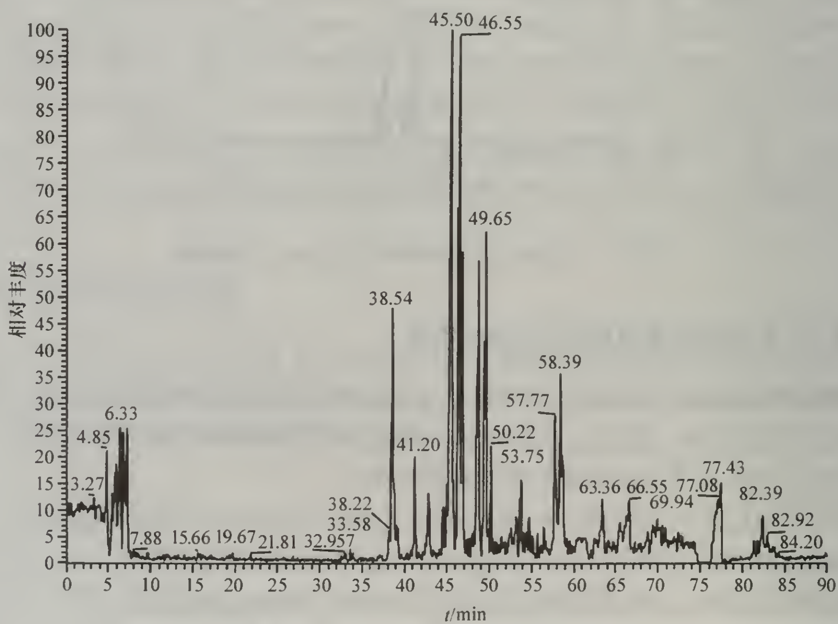


图 17.6　ICAT 试剂标记后的 β-lactoglobulin 胰蛋白酶酶解肽质谱离子流图

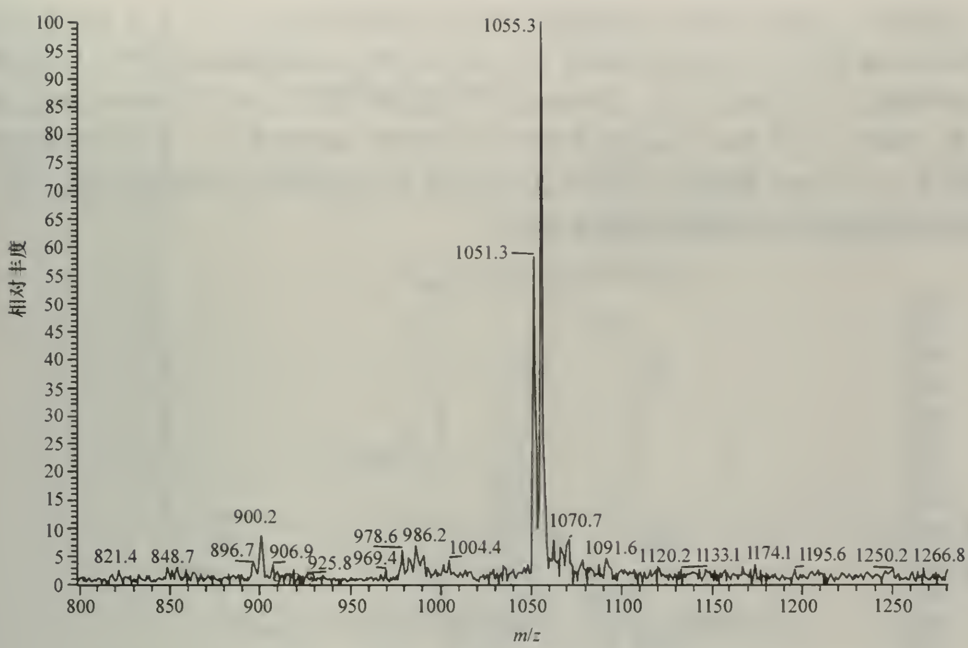


图 17.7 β-lactoglobulin 中被 ICAT 试剂标记肽 LSFNPTQLEEQCHI 的质谱图

图中两个峰分别为 D_0标记肽和 D_8标记肽

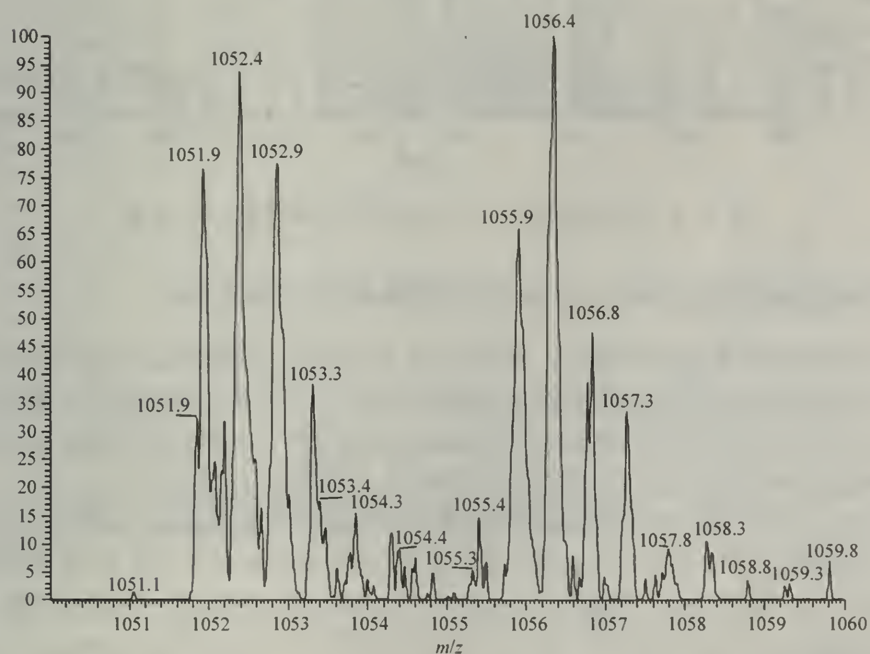


图 17.8 D_0和 D_8试剂标记肽 LSFNPTQLEEQCHI 的 ZoomScan 显示该肽段在质谱中为双电荷形式

从图 17.7 中可看到，D_0标记的 β-lactoglobulin 和 D_8标记的 β-lactoglobulin 质量比为 1∶2的图谱，实验数据得到的比例是 1∶1.8，这说明了 ICAT 方法定量的可靠性。

但要对这一肽段进行确认，还应该进行 MS/MS 的碎片分析。图 17.9 显示的是 D_8标记的相对分子质量为 2 108.6 的肽段的 MS/MS 碎片图。肽键断裂后形成两个系列的碎片，分别被称为 b 系列和 y 系列。图中标出了匹配的碎片质量，丰度较高的碎片基本上都被匹配。因此，可以认为这一肽段就是图中所标的序列，也就是表 17.1 中所示的相对分子质量为 1 658.7843 的肽段 LSFNPTQLEEQCHI 在 D_8试剂标记后得到的产物。另一 D_0标记的肽段的情况大致相似(结果未显示)。

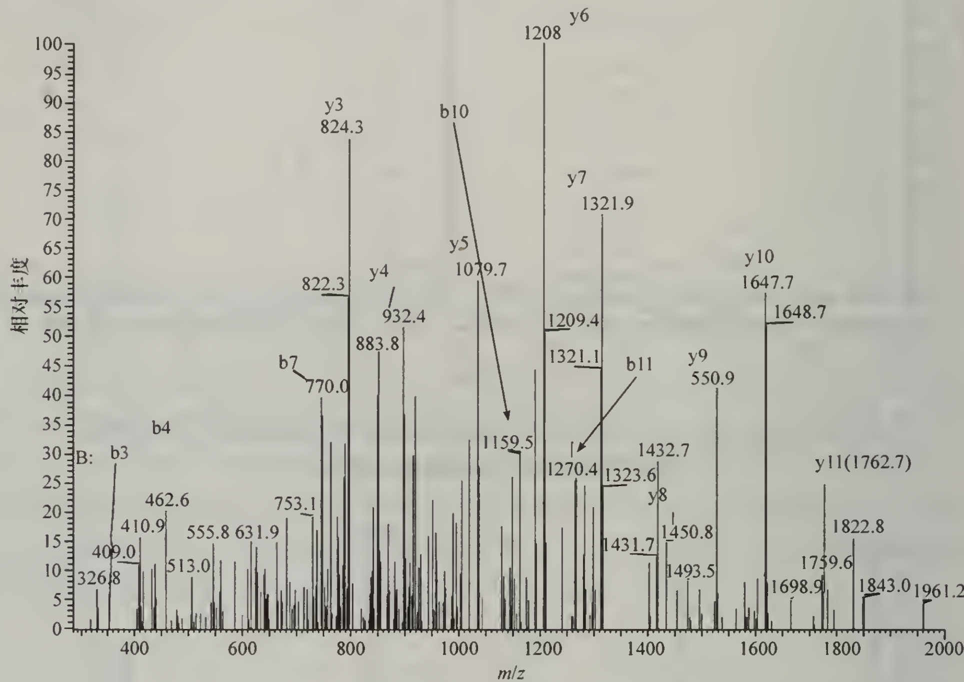


图 17.9　D_8试剂标记肽 LSFNPTQLEEQCHI 的 MS/MS 图谱

3. 细胞抽提总蛋白质的 ICAT 标记后的多维色谱质谱联用分析

由于多维色谱强大的分离能力，所以 ICAT 方法在蛋白质组研究中的应用时，没有对标记后的肽段进行亲和色谱纯化，而是直接进行分析。在对钩端螺旋体进行蛋白质组研究中，使用 ICAT 方法对不同菌株的蛋白质表达进行分析。分析的两种菌体分别为有毒株(V)和无毒株(NV)，其中有毒株用 D_0试剂标记，无毒株用 D_8试剂标记。

图 17.10 所示为分析的一种未知功能蛋白质在有毒株和无毒株之间的表达差异，图谱为同一段肽(LCTHSESVLK)的两个标记形式，因为离子带有两个电荷，所以质荷比相差 4。图 17.11 所示为这一段肽的 D_0标记形式的 MS/MS 图谱，图中显示碎片峰中大多数都和该肽段的 b 离子和 y 离子一致，并可以清楚地显示出 D_0标记的特征峰 m/z 284.1，充分证明鉴定的准确性，且为 ICAT 试剂标记的肽。图 17.8 中的两种标记肽段的离子强度比值为 1∶0.46，说明该蛋白质在有毒株中的表达高于无毒株约 2 倍。这是一个未知蛋白质，在有毒株中的高表达意味着该蛋白质可能与钩端螺旋体的致病性或毒力有关。

图 17.12 所示为电子转移黄素蛋白 α 亚基(electron transfer flavoprotein alpha-sub-

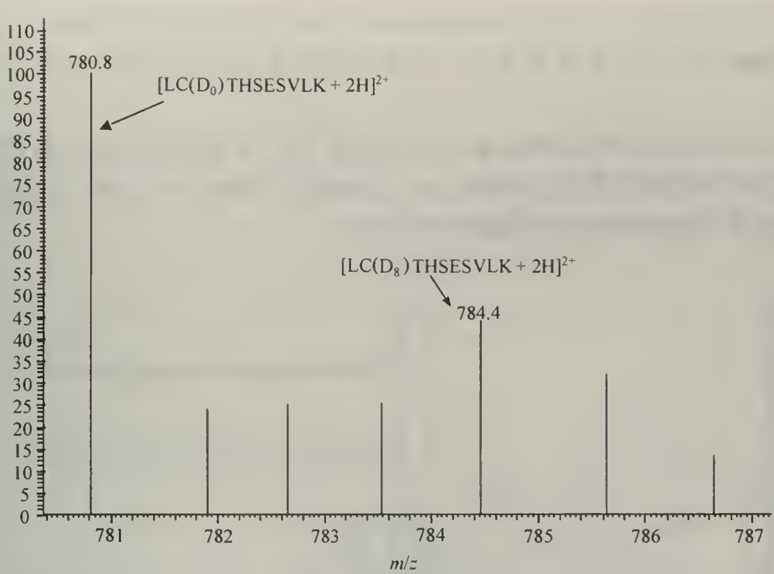


图 17.10　D_0和 D_8试剂标记肽 LCTHSESVLK 的质谱图

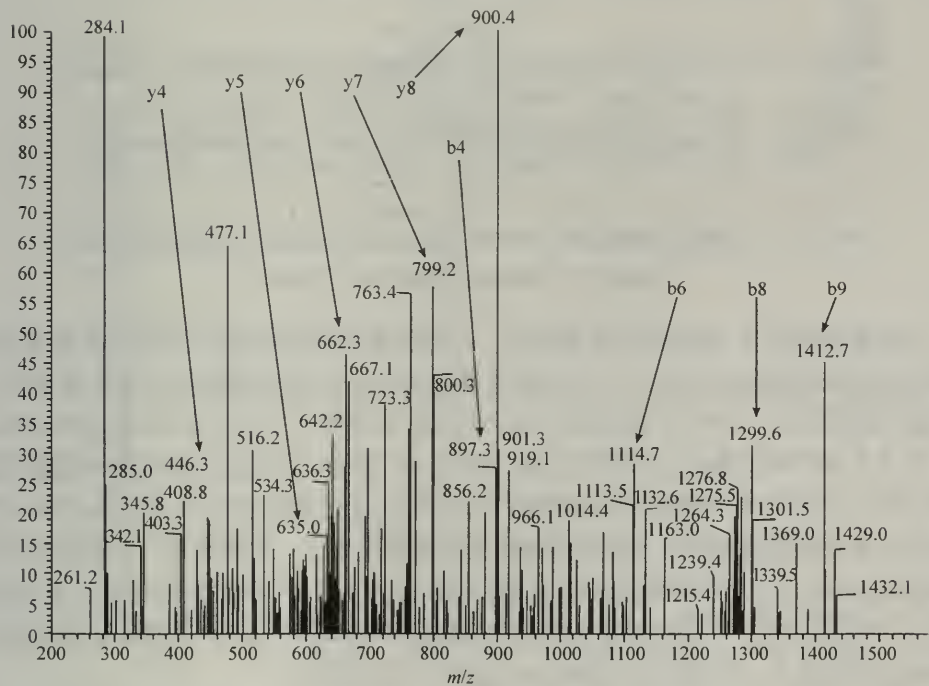


图 17.11　D_0试剂标记肽 LCTHSESVLK 的 MS/MS 图谱

unit)在两个菌种之间的表达差异，该肽段中有两个半胱氨酸，离子带有三个电荷，所以质荷比相差 5。两者的离子强度比值为 1∶0.47。这个差异与双向凝胶电泳所分析的结果一致(结果未显示)。图 17.13 和图 17.14 为该肽段的两种标记形式的 MS/MS 图谱，图中显示碎片峰中大多数都和该肽段的 b 离子和 y 离子一致，充分说明鉴定的准确性。该肽

段的序列为 TVSPNCYIACGISGAIQHLAGMGSSK。该蛋白质在有毒株中的表达高于无毒株。目前,关于电子转移黄素蛋白 α 亚基与钩端螺旋体的致病性或毒力相关还未见报道。

ICAT 方法是近几年来发展的蛋白质组定量方法,实验证明在蛋白质组研究中有良好的应用前景,其分析速度快,所耗样品少,完全脱离了凝胶分离过程,与质谱技术兼容性好,大大促进了蛋白质组研究大规模定量的应用。

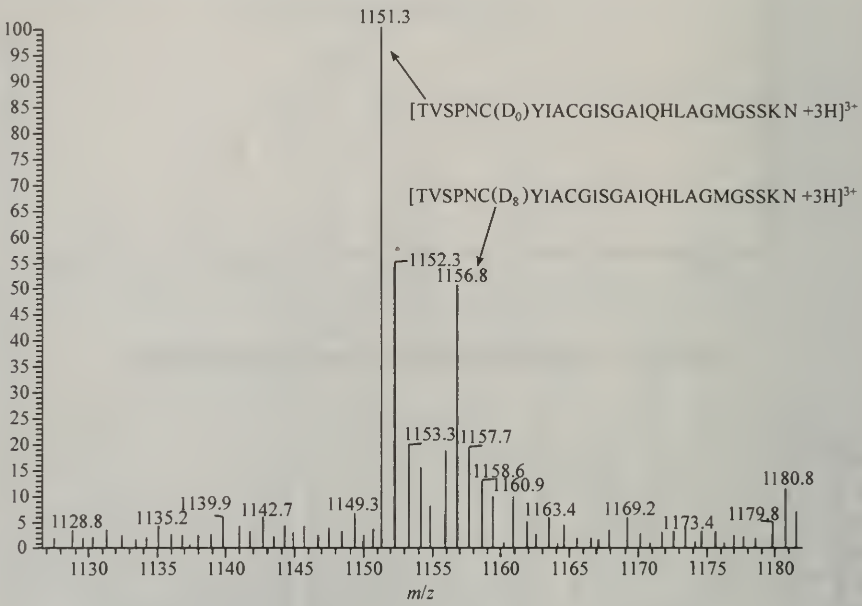


图 17.12 D_0和 D_8试剂标记肽 TVSPNCYIACGISGAIQHLAGMGSSK 的质谱图

肽段有两个半胱氨酸,在质谱中带 3 个电荷

对定量技术而言,准确性是最重要的。影响准确性的因素很多,实现样品制备和总蛋白质定量的精确是基本保证。另外还要尽量保持样品在还原和标记过程中条件的一致性,减少实验误差。另外,目前使用的 ICAT 方法本身也有一些弱点,需在结果分析时注意。本文介绍的是用氢(D_0)和氘(D_8)标记的 ICAT 试剂。标记上这种试剂的肽段在色谱行为上有一定的区别,出峰时间有细微差异,通常 D_8标记的肽会比 D_0标记的肽较早被洗脱,因此,在液相色谱质谱联用使给质谱检测和定量带来了一定的困难。一般 ICAT 定量的误差在 20%以内,但如果样品成分复杂,定量误差会增大,并往往需要手工检查结果。目前,一种新型的用^{13}C 标记的 ICAT 试剂已经商品化,这种 ICAT 试剂标记的肽段在保留时间上没有差异, 定量将更为准确(本实验室未发表结果)。

ICAT 方法的另一个问题是,试剂中的 Biotin 基团相对于肽段是相当大的,在做二级质谱碎片分析时,巨大的修饰对肽段碎片的形成不利,在大规模分析时会造成蛋白质鉴定的分数偏低。进一步的实验中已经将 Biotin 修饰部分在亲和纯化后去除,提高了肽段碎片谱的质量。

ICAT 技术定量依赖于 ICAT 试剂特异性地结合于半胱氨酸残基的游离巯基上,但是在细胞总蛋白质中(以酵母为例),约 8% 的蛋白质是不含有半胱氨酸的,就无法用

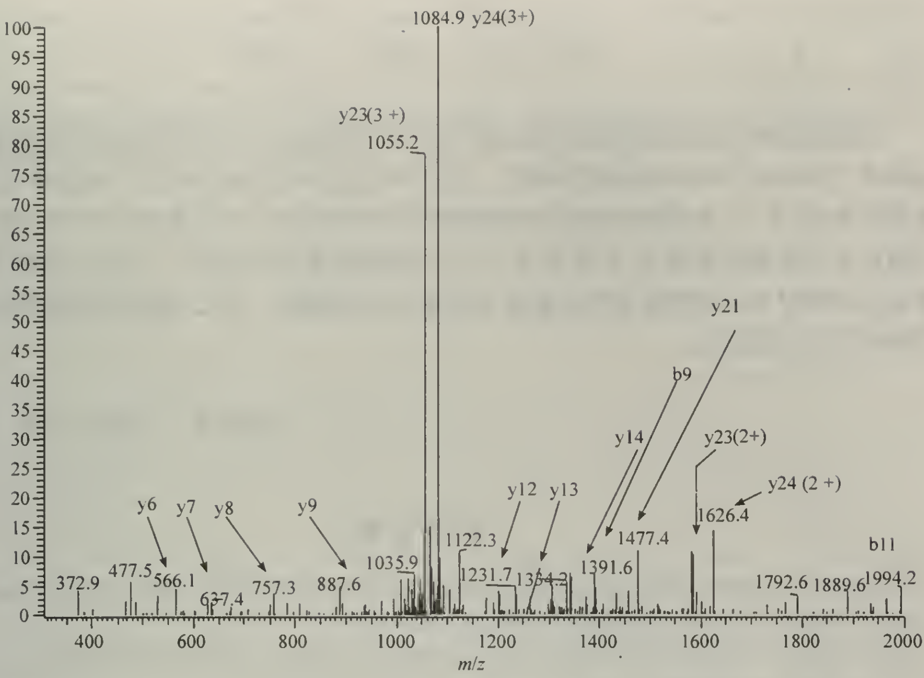


图 17.13　D_0试剂标记肽 TVSPNCYIACGISGAIQHLAGMGSSK 的 MS/MS 图谱

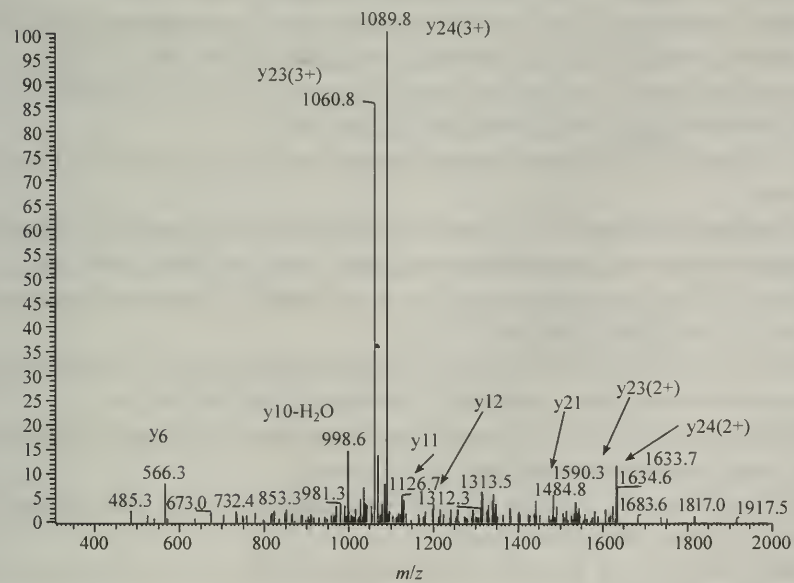


图 17.14　D_8试剂标记肽 TVSPNCYIACGISGAIQHLAGMGSSK 的 MS/MS 图谱

ICAT定量。所以，目前也有发展了其他方法如标记在赖氨酸、精氨酸的同位素氨基酸掺入法等。但就标记特异性和使用简便程度而言，ICAT 技术的优势还是明显的。就定量蛋白质组学研究的发展方向而言，以 ICAT 技术为代表的稳定同位素标记技术将发挥越来越大的作用。

第五节 展 望

蛋白质组研究与基因组研究相比最大的不同和难点之一就是定量,而蛋白质表达量的差异又是影响生物功能的重要因素。蛋白质组定量技术现在还处于起步阶段,也面临着很多难题:首先,对低丰度蛋白质检测的困难显然阻碍了对这些蛋白质的定量。其次,当蛋白质表达量差异很小,如在50%以下时,精确定量成为瓶颈。另外,生物体系中蛋白质表达的瞬时变化的捕捉是样品制备中需要关注的问题。总之,蛋白质组定量技术的研究和应用还任重道远。

(曹兴军 阮宏强 曾 嵘)

参考文献

1 Wayne FP. A thousand points of light: the application of fluorescene detection technologies to two-dimensional gel electrophoresis and proteomics. Electrophoresis,2000,21:1123~1144

2 Oda Y,Huang K,Cross F *et al*. Accurate quantitation of protein expression and site-specific phosphorylation. Proc Natl Acad Sci USA,1999,96:6591~6596

3 Conrads TP,Alving K,Veenstra TD *et al*. Quantitative analysis of bacterial and mammalian proteomes using a combination of cysteine affinity tags and ^{15}N-metabolic labeling. Anal Chem,2001,73:2132~2139

4 Ong SE,Blagoev B,Kratchmarova I *et al*. Stable isotope labeling by amino acida in cell culture,SILAC,as a simple and accurate approach to expression proteomics. Mol Cell Proteomics,2002,1.5:376~386

5 Gu S,Pan S,Bradbury EM *et al*. Precise peptide sequencing and protein quantification in the human proteome throμgh in vivo lysine-specific mass tagging. J Am Soc Mass Spectrom,2003,14:1~7

6 Chakraborty A,Regnier FE. Global internal standard technology for comparative proteomics. J Chromatogr A,2002,949:173~184

7 Mü nchbach M,Quadroni M,Miotto G *et al*. Quantitation and facilitated de nevo sequencing of proteins by isotopic N-terminal labeling of peptides with a fragmentation-directing moiety. Anal Chem,2000,72:4047~4057

8 Cagney G,Emili A. *De novo* peptide sequencing and quantitative profiling of complex protein mixtures using mass-coded abundance tagging. Nat Biotechnol,2002,20:163~170

9 Wang S,Zhang X,Regnier FE. Quantitative proteomics strategy involving the selection of peptides containing both cysteine and histidine from tryptic digests of cell lysates. J Chromatogr A,2002,949:153~162

10 Goodlett DR,Keller A,Watts JD *et al*. Differential stable isotope labeling of peptides for quantitation and de novo derivation. Rapid Commun Mass Spectrom,2001,15:1214~1221

11 Reynolds KJ,Yao X,Fenselau C. Proteolytic ^{18}O labeling for comparative proteomics: model studies with two serotypes of adenovirus. Anal Chem,2001,73:2836~2842

12 Gygi SP,Rist B,Gerber SA *et al*. Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. Nat Biotechnol,1999,17:994~999

13 Griffin TJ,Gygi SP,Ideker T *et al*. Complementary profiling of gene expression at the transcriptome and proteome levels in *Saccharomyces cerevisiae*. Mol Cell Proteomics,2002,1:323~333

14 Han DK,Eng J,Zhou H *et al* . Quantitative profiling of differentiation-induced microsomal proteins using isotope-coded affinity tags and mass spectrometry. Nat Biotechnol,2002,19:946~951

15 Ideker T,Thorsson V,Ranish JA *et al*. Integrated genomic and proteomic analyses of a systematically perturbed metabolic network. Science,2001,292:929~934

16 Zhou H, Ranish JA, Watts JD *et al*. Quantitative proteome analysis by solid-phase isotope tagging and mass spectrometry. Nat Biotechnol, 2002, 19:512～515

17 Timothy JG, Aebersold R. Advance in proteome analysis by mass spectrometry. J Biol Chem, 2001, 276:45497～45500

18 Goshe MB, Conrads TP, Panisko EA *et al*. Phosphoprotein isotope-coded affinity tag approach for isolating and quantitating phosphopeptides in proteome-wide analysis. Anal Chem, 2001, 73:2578～2586

第十八章　蛋白质组研究中的翻译后修饰分析

第一节　磷酸化蛋白质组研究

一、介　　绍

蛋白质磷酸化研究是一项大事业。在生命现象的许多关键调节机制中，蛋白质的磷酸化是最重要的翻译后修饰。20 世纪 50 年代，发现磷酸化酶 a 和磷酸化酶 b 原来是同一种酶的磷酸化和脱磷酸化形式。从那时开始，生物学家开始将蛋白质的磷酸化看作是一种动态的生物调节过程[1]。在原核生物中，磷酸化位点主要位于组氨酸、谷氨酸和天冬氨酸残基，但真核生物中的磷酸化主要发生在丝氨酸、苏氨酸和酪氨酸残基[2]。自从 20 世纪 50 年代以后，磷酸化蛋白质的研究一直受到生物学家的广泛重视，重大的研究成果屡见不鲜。但是，传统的磷酸化分析依赖于放射性同位素标记、Edman 降解以及薄层层析等方法[2]，这些方法冗长，需要高超的实验技巧和较多的蛋白质，而且对操作者存在放射性危险[3]。

20 世纪 80 年代，蛋白质分析领域发生了一场重要的技术革命。“电喷雾”——第一种可以运用于生物大分子的质谱电离方式诞生了，随后，基质辅助激光解析电离也应用于蛋白质的分析，质谱分析成为蛋白质分析领域的中坚力量。当然，作为蛋白质分析“新宠”的质谱技术，也被大量用于磷酸化蛋白质的检测。

90 年代后，质谱技术的发展迎来了历史上最好的阶段，世界各国投入大量精力，开发质谱技术。各种类型的质谱纷纷出现：如 MALDI-TOF、MALDI-QTOF、MALDI-TOF-TOF、MALDI-ION-TRAP、ESI-ION-TRAP、ESI-Q-TOF、三级四极杆质谱和即将成为质谱技术先锋的 FTICR。本文将就使用现代质谱技术对真核生物蛋白质磷酸化的分析检测及磷酸化蛋白质组研究的策略进行综述。

（一）概　　况

蛋白质磷酸化在真核生物中是非常重要而且是非常普遍的现象。据估计，在任一给定时刻，细胞内约有 1/3 的蛋白质存在磷酸化[4]。在真核生物常见的三种磷酸化形式中，丝氨酸磷酸化最多，苏氨酸磷酸化其次，而酪氨酸磷酸化最少[5]，三者的比例是 1 800∶200∶1。磷酸化对于蛋白质而言是非均一性的，大多数磷酸化蛋白质都有多个磷酸化位点，但这并不意味着某一蛋白质分子的所有潜在的磷酸化位点都是磷酸化的[6]。人类的基因组中有 2% 的序列用来编码磷酸化蛋白质，主要是蛋白质激酶和磷酸酯酶，其中丝氨酸和苏氨酸激酶的数量是酪氨酸激酶的 4 倍，而它们的磷酸酯酶的数量基本相同[3]。

要对蛋白质磷酸化进行分析，首先要对蛋白质磷酸化的化学有所了解。丝氨酸和苏氨酸磷酸化是不稳定的，在碱性条件下，磷酸基团会脱去，发生 β 消除（图 18.1）。酪氨酸的磷酸化由于磷酸基团连在苯环上，不会发生 β 消除，相对是较稳定的。丝氨酸和苏氨酸

残基上磷酸基团脱去后的质量变化与酪氨酸是不同的(图 18.2),酪氨酸因为无法进行β消除反应。所以,在质量变化上比丝氨酸和苏氨酸少一步,减少的分子质量是 80Da,而不是 98Da,这些细节的质量变化是磷酸化质谱检测中所必须注意的。

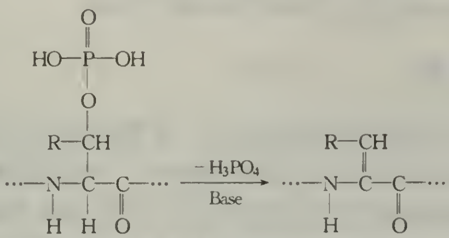
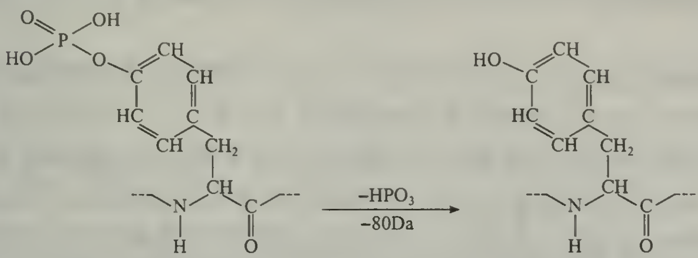
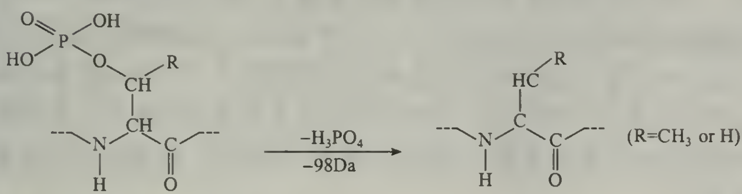


图 18.1　丝氨酸和苏氨酸磷酸化的不稳定性($R=H、CH_3$)

磷酸基团在碱性环境下易于脱去,并伴随着β消除反应,再脱去一分子水



磷酸化酪氨酸残基　　　　酪氨酸残基



磷酸化丝氨酸和磷酸氏苏氨酸残基

图 18.2　氨基酸残基上的磷酸基团脱落时发生的结构变化和质量变化

分析蛋白质磷酸化是困难的,尽管现代质谱技术的发展非常迅猛,但是磷酸化面临的难点仍然存在。具体分析有以下几点。首先,磷酸化蛋白质在细胞内的蛋白质中是相对较低丰度的。第二,即使我们找到一种磷酸化蛋白质,也不能排除有该蛋白质的其他磷酸化形式存在,在体内磷酸化与其说是一座“山”(静态),不如说是一条“河”(动态)。第三,细胞内有很多磷酸酯酶,在样品处理时,这些酶很容易将磷酸基团脱掉。因此,在处理样品时,加入磷酸酯酶抑制剂是非常必要的[6]。第四,磷酸化蛋白质酶解后的磷酸化肽段,因为其化学性质的负电性,在质谱技术中面临着难以质子化的困难。此外,磷酸化肽段在质谱图谱中还往往受到非磷酸化的肽段信号的强烈抑制,这一点在 MALDI 源的质谱中尤为明显[7,8]。

（二）富集的策略

由于磷酸化蛋白质的丰度低，而且磷酸化肽段分析检测存在困难，使得富集在磷酸化蛋白质组的研究中显得十分重要。富集的策略应用在磷酸化蛋白质组研究中，有两种不同的应用阶段，包括富集磷酸化蛋白质和富集磷酸化肽段。

1. 磷酸化蛋白质的富集

这一类富集策略的目标是磷酸化的蛋白质。富集这些蛋白质的主要手段是使用磷酸化蛋白的抗体。磷酸化蛋白质的抗体主要分成两种：一种是丝氨酸和苏氨酸磷酸化蛋白质的抗体，另一种是酪氨酸磷酸化蛋白质的抗体。其中，酪氨酸磷酸化抗体的应用较早，已有专一性强、效率高的抗体存在，因此也较早用于免疫共沉淀[9]。最近，已经有可用于丝氨酸和苏氨酸磷酸化的抗体面世，这些抗体不但可以用于 Western 印迹检测，还可以用于免疫共沉淀纯化磷酸化蛋白质，将大大方便免疫共沉淀在磷酸化蛋白质的富集中的应用[6]。

同样是富集磷酸化蛋白质，也有采用化学方法进行修饰，改变磷酸基团的化学性质，然后特异性分离纯化，并可以定量分析这些磷酸化蛋白质[10~12]。这些方法无一例外地涉及到化学修饰的效率问题，在试验中蛋白质的损失是难免的，这也间接影响了此类方法的应用。其中一种化学修饰的基础是利用磷酸基团脱去后的 β 消除反应，通过加成反应，共价连接新的化学修饰基团。β 消除反应后生成的脱氢氨基酸直接由质谱检测（MS/MS），或用 EDT（ethanedithiol）作为亲核试剂，提供一个新的巯基以便连接一个生物素的亲和标签，为下一步的亲和纯化和同位素标记定量（ICAT）提供条件[11,13]。肽段中的半胱氨酸和甲硫氨酸会和 EDT 发生副反应，因此首先要对这两种残基进行强烈氧化，使其丧失活性。Goshe 等人使用同位素标记的 EDT，这提供了一个巧妙的解决磷酸化肽段定量问题的方案[12]（如图 18.3 所示）。这种化学方法的优点是所有的反应都在同一个反应体系内进行，没有很多的反应步骤，反应的产率有一定保证；其缺点是对于酪氨酸的磷酸化是无能为力的。

第二种化学修饰是第一种方法的合理补充，它可以应用于酪氨酸磷酸化的检测。这种方法在保护肽段的氨基和羧基端的基础上，对磷酸基团进行化学修饰，然后用特异化学表面的微球纯化，很明显，这一方法的化学基础复杂，其核心是对磷酸基团的修饰，以便和有特异化学表面的玻璃珠结合纯化。这一方法的缺点是步骤太多，产率相对较低，这一点连作者本人都不否认[10]。详细化学过程见图 18.4。

2. 磷酸化肽段的富集策略

磷酸化蛋白质组研究中应用的抗体只是对磷酸化蛋白质有效。对磷酸化肽段进行富集使用较多的是固定金属螯合亲和层析（immobilized metal affinity chromatography, IMAC）。

磷酸化酪氨酸残基由于结构的制约，不能像磷酸化丝氨酸和苏氨酸那样发生 β 消除反应，因此分子质量数的变化只有 80Da，而一般磷酸化丝氨酸和苏氨酸的质量数变化有 98Da。

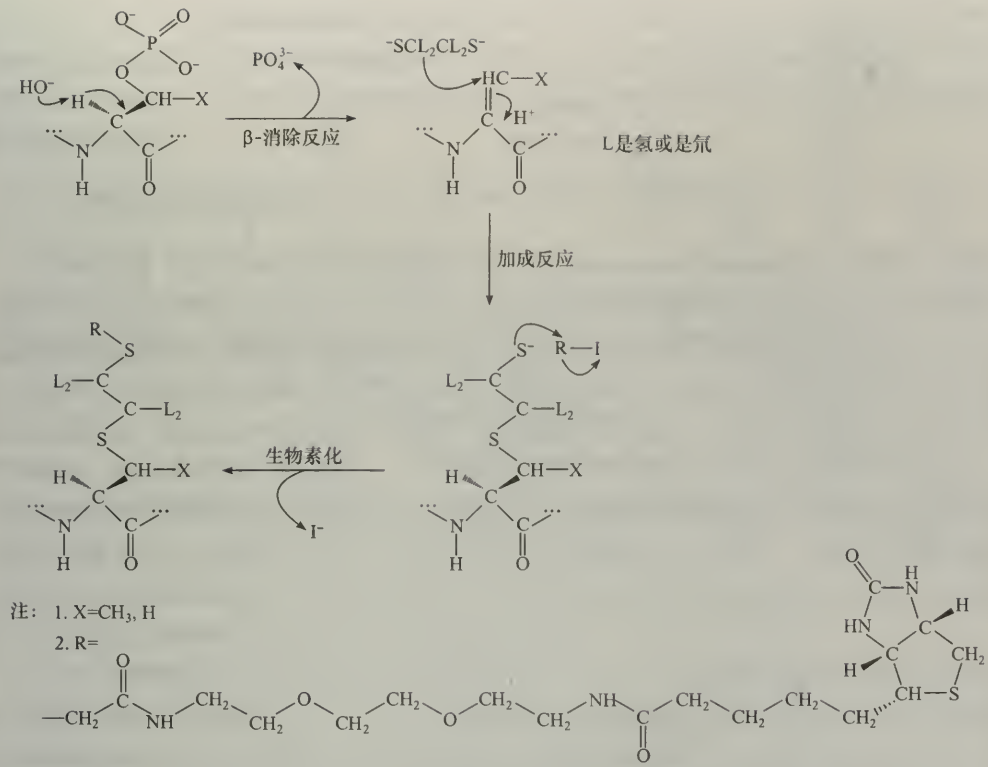


图 18.3　通过同位素标记的化学修饰进行丝氨酸和苏氨酸磷酸化肽段的富集和定量

这项技术从 20 世纪 90 年代开始应用于磷酸化肽段的检测，在 online 和 offline 两个形式上都得到了满意的结果[14~19]。IMAC 技术是用于富集磷酸化的肽段，对磷酸化蛋白质的富集是无效的。IMAC 技术使用的金属离子主要是 Fe^{3+} 和 Ga^{3+}，据文献[16]报道 Ga^{3+} 的富集效果比 Fe^{3+} 要好，但是本实验室的实践和后续的文献报道[20]显示，Fe^{3+} 的效果是可以接受的，而且由于 Fe^{3+} 的价格便宜和易于操作的优点，比 Ga^{3+} 的应用要广泛些。

蛋白质组学的开展，需要有大规模、自动化的技术。对 IMAC 技术的要求也是如此，尽管有报道，在 offline 条件下的磷酸化肽段纯化有良好的结果，但是从长远来看，IMAC 参与其中的多维色谱技术的发展是必须的。在这一方面，已经显著进展。Zarling 等人在 2000 年报道使用内径 100μm 的 IMAC 柱，与两根反相柱联用形成多维色谱，取得了良好的试验效果[21]。之后不久，在 *Nature Biotechnology* 杂志上，就报道了更出色的工作[20]，在使用基本相同的技术手段的情况下，在酵母的全细胞裂解液中检测到了超过1 000个磷酸化肽段，并分析了其中的一部分。

IMAC 柱富集磷酸化肽段的原理，是利用磷酸基团的负电性与 IMAC 柱上的带正电的金属离子结合，从而起到纯化的作用。IMAC 技术对丝氨酸、苏氨酸和酪氨酸磷酸化都有效，但是 IMAC 柱上的正电荷位点同样会与肽段中的天冬氨酸、谷氨酸和组氨酸残基结合，从而特异性和非特异性肽段同时富集。所以，IMAC 方法面临的最主要的缺点是特异性不强。在上述同一篇文献中，对磷酸化肽段中的酸性残基进行了预先的甲基酯化，封闭了上述氨基酸的酸性侧链，从而大大提高了 IMAC 方法的特异性[20]。这一方法弥补了

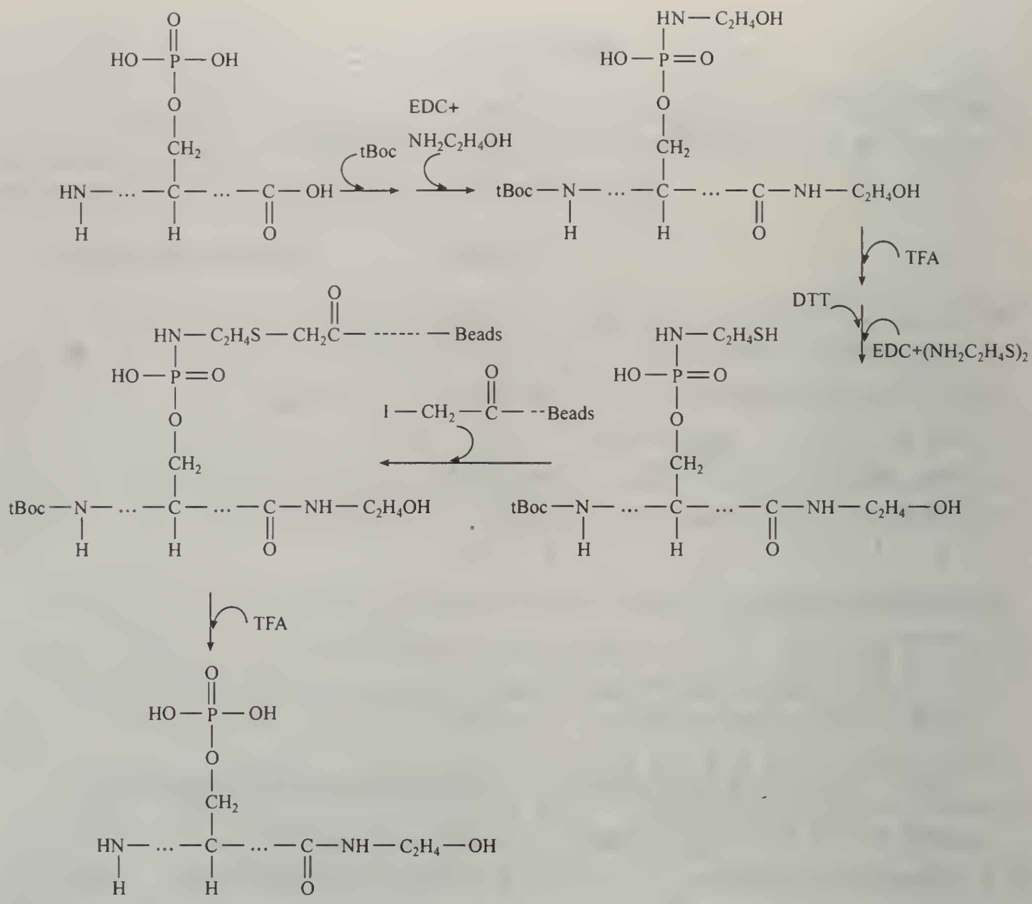


图 18.4　化学修饰后进行磷酸化蛋白质的纯化(碳化二乙胺反应)

IMAC 技术最大的缺陷,为大规模磷酸化蛋白质组学的开展展现了美好的前景。

磷酸化蛋白质和肽段的富集对于它们的检测是重要的,进行磷酸化蛋白质组研究,第一步要做的就是富集,以上介绍的这两方面的富集策略,各有优缺点。笔者认为,以抗体为基础的第一种策略有效但昂贵,现阶段酪氨酸磷酸化蛋白质的抗体比丝氨酸和苏氨酸的要有效些,对于大多数国内实验室来说,受财力所限使用时会有困难;而以色谱方法为基础的 IMAC 方法,同样是有效的,甚至于在有化学修饰辅助下是更有效的[20],重要的是其试验花费和前者相比是微不足道的,因此会有更好的应用。本实验室的实践证明在国内现有的条件下,使用相对简单的 IMAC 色谱柱对磷酸化肽段的富集也能得到满意的结果。

(三)磷酸化肽段的质谱检测

当代质谱技术沿着两条主线发展,其一是离子源,其二是检测器,这两个方向上几乎包括了所有新的技术思路。就离子源而言,自从 20 世纪八九十年代发明 ESI 和 MALDI 后,鲜有质的突破。时至今日,绝大多数生物质谱的离子化方式仍然是这两种。而就质量分析器而言,种类倒是颇多,有经典的三级四极杆、离子阱和飞行时间检测器,也有稍后的

Qq-TOF、TOF-TOF 和即将成为热门的 FTICR。近几年来，质谱技术的发展主要集中在各种离子源和检测器、检测器和检测器的“杂交”上，产生了多种新颖的生物质谱，应用于生物学各个领域。以下将从生物质谱的这些技术特征角度，分别介绍应用于蛋白质组磷酸化研究的情况。

1. 以 MALDI 为离子源

以 MALDI 为离子源的生物质谱通过 PMF(peptide mass fingerprinting)技术，在鉴定蛋白质的工作中取得了很好的结果，但将 MALDI 离子源的质谱应用于蛋白质组磷酸化研究就遇到了问题。如前所述，磷酸化肽段是带有负电性的，而生物质谱常用正离子模式，它的离子化就显得不佳。更糟糕的是进行磷酸化研究时，磷酸化肽段常常和非磷酸化的肽段混合在一起，在 MALDI 源的情况下，磷酸化肽段的信号会被非磷酸化肽段的质谱信号抑制。在未经富集的情况下，一个磷酸化蛋白质的磷酸化肽段在 MALDI 源时，经常会淹没在其他非磷酸化肽段的信号中，连信号都没有，分析研究磷酸化的位置、数量和变化就成了奢望。此外，因为以 MALDI 为离子源的质谱没有分离能力，所以蛋白质必须先进行纯化，在蛋白质组学中，现在应用最广的纯化方法还是双向电泳和 HPLC。因此，MALDI 质谱分析磷酸化时，就会有些必要的前处理。

首先，磷酸酯酶被应用到磷酸化肽段的检测中[18,21,22]。磷酸酯酶处理蛋白质的前后，磷酸化肽段的质量数会有变化，这为磷酸化肽段的检测提供了前提条件；更重要的是，很多磷酸化肽段在脱去磷酸基团后，质谱信号会有很大提高[18]。这样比较前后图谱，寻找质量数变化 80 或 98Da 和强度增大的信号就很可能是磷酸化肽段。在磷酸酯酶反应时，丝氨酸和苏氨酸磷酸化肽段的质量变化可能是 80Da，也可能是 98Da，后者的发生是因为在磷酸基团脱落后发生了 β 消除反应，酪氨酸磷酸化肽段只有 80Da 的质量数变化。磷酸酯酶的另一种应用方式是固定化磷酸酯酶，使脱磷酸作用发生在亲和柱上，这提高了酶反应的效率和方法的自动化程度[23]。

其次，以 MALDI 为离子源的质谱检测磷酸化肽段时，对磷酸化肽段的富集显得更为重要。这其中，IMAC 技术与之合用较为多见[14~16,18]。磷酸化肽段经过富集，在MALDI 质谱上的检测信号会有较大的提高，对于这一点，本实验室对 β 酪蛋白的酶解产物进行了 IMAC 的富集，其磷酸化肽段的信号得到明显增强。

MALDI 离子源质谱较早使用的是 TOF 检测器，受其原理所限，MALDI-TOF 质谱的 PSD 碎片分析非常困难。为了克服 MALDI 质谱的这一重要缺陷，近年来，发展了多种以 MALDI 为离子源的“杂交”质谱，如 MALDI-ION-TARP、MALDI-Qq-TOF 和 MALDI-TOF-TOF 等等。这些技术的出发点都是为了在 MALDI 离子源的情况下，进行肽段的碎片分析。MALDI-IONTRAP 质谱是其中较早出现的一种[24,25]，思路是将 MALDI 源产生的离子选择性的用离子阱俘获，并作 CID(collision-induced dissociation)分析，得到肽段碎片信息，以便更好地分析鉴定蛋白质。在选择作 CID 的离子时，主要选择有成对出现的质量数相差 98Da 的质谱信号。MALDI-ION-TRAP 技术很早就被认为是有前景的[7]。当时此项质谱技术还停留在实验室阶段，而现在这种质谱已经被成功地商品化，并很快被应用于磷酸化肽段的检测上来[26]。Chait 等人的工作说明，在这种质谱条件下，丝氨酸、苏氨酸和酪氨酸磷酸化肽段在 CID 分析时，都会形成失去 98Da 的碎片峰[25]。丝氨酸和

苏氨酸磷酸化肽段 CID 时失去的 98Da 的质量数一般被认为是伴随着 β 消除反应而发生的(也有人认为是 *cis*-1,2-elimination,而不是 β-elimination[27]),丝氨酸和苏氨酸磷酸化肽段磷酸基团的脱落是常见的现象,而且与肽段的氨基酸构成没有关系。而酪氨酸磷酸化肽段在 CID 时的质量数变化也是 98Da,这很有意义。现在认为这里发生了两步反应,第一是 HPO_3的失去,第二是一分子水的失去。同时,酪氨酸磷酸化肽段的磷酸基团较前两者不易脱落(这些脱落与肽段的氨基酸组成有关,在肽段中有碱性氨基酸赖氨酸和精氨酸存在时有利于磷酸基团的脱落),因此,在 CID 时能看到带有磷酸的肽段碎片,这一点可能对磷酸化位点的判断有帮助。MALDI-IONTRAP 也有令人遗憾的缺点,因为 MALDI 离子源情况下,产生的离子大多是单电荷的,而单电荷的离子在离子阱内作 CID 时信号显得不太强烈,产生的碎片峰的强度也不太理想,这不能不说是该项技术的一个隐忧。

2. 以 ESI 为离子源

可以毫不夸张地说,ESI(电喷雾)离子源质谱在磷酸化肽段的检测中,有着比MALDI离子源质谱更广泛的应用和更好的前景。在大多数情况下,ESI 离子源质谱都会配以一个 HPLC 系统作为分离的工具,这使其可以分析较为复杂的系统,而不像很多 MALDI 离子源质谱只能分析相对较纯的样品或者是不太复杂的肽段混合物。近年来,质谱技术与毛细管 HPLC 技术相结合,而且越来越多的实验室开始使用 nano-spray 的电喷雾技术,大大的提高了 ESI 技术的灵敏度和应用范围[28]。同时,也是由于 ESI 离子源与色谱技术的良好兼容性,多种现代色谱技术参与到其中,与 ESI 质谱联用,大大的推动了质谱本身的发展。比如,多维色谱技术、IMAC 技术、微柱 HPLC 以及填料技术等等都越来越多的在质谱技术中发挥作用。可以说,以电喷雾技术为离子源的质谱是蛋白质组学研究的主力,同时也是磷酸化蛋白质组研究的重要武器。这方面,磷酸化肽段的分析方法是丰富多彩的,其核心是围绕着肽段磷酸化的检测。

(1)前离子扫描(Precursor ion scan)

前体检测的思路是,利用磷酸基团的化学特性,检测它在各种条件下所生成的有特点的离子。这其中有带负电的 PO_4^{3-},也有带正电的 immonium 离子。前离子扫描这项方法的离子源都是电喷雾的,而质量分析器主要是串联质谱,如三级四级杆和 Qq-TOF 等。

前体离子中较早被用来做前离子扫描的是 PO_4^{3-}(m/z 79)。PO_4^{3-}是在低能量的 CID 时,从磷酸化肽段上脱落的小分子离子,将它作为标识分子进行检测,然后在串联质谱中对它的母离子进行分析。在使用三级四级杆分析器时,经典的方法是,第一级四级杆分析器进行全质量数范围扫描,第二级四级杆分析器进行 CID 碎片分析,而第三级四级杆分析器则专门扫描 PO_4^{3-}(m/z 79)[29]。在找到这个特征性的小分子的时刻,质谱上记录下的肽段和肽段 CID 的信息就是磷酸化肽段的的信息。之所以能够对 PO_4^{3-}进行前离子扫描,作为磷酸化肽段的特征性标识,原因是现代生物质谱对这样的小分子有着很高的灵敏度和准确度。对 PO_4^{3-}的前离子扫描,在多项研究中得到了应用[30,31],但是这里有个矛盾的问题,如果没有 HPLC 进行分离,前离子扫描的分析是不可想像的,但是,HPLC 的酸性环境对质谱的负离子模式的灵敏度是非常不利的。尽管有人尝试使用多维色谱来解决所遇到的困难[32],但是这种努力显然遇到了很大的技术阻力,没有获得如意的前景。

前体离子扫描的发展是很快的,面对负离子模式下一些难以逾越的困难,另一些正离

子模式下的前体离子扫描方法被发展起来[33-35]。较早一些的是扫描酪氨酸磷酸化的 immonium(pY) 离子(m/z 216.043),最早 Hoffmann R 等人首先描述到 immonium 离子是酪氨酸磷酸化肽段的特征性的碎片离子[36]。2001 年,丹麦学者将这种离子的前离子扫描付诸实践[35]。他们使用的质量分析器是 Q-TOF,之所以要使用这种分析器,是因为其精确度比三级四极杆要高,而要在一系列的质量数相差很小的碎片离子中准确的找到 immonium 离子,需要相当高的精确性。他们的工作显示,这种方法在标准品上可以达到 1 飞摩尔的水平,而在胶内蛋白质也只要 100 飞摩尔就足够了。这方法可以作为研究酪氨酸磷酸化时的不错的选择。另一些正离子模式的方法,涉及到对残基上的磷酸基团的化学修饰[33],在磷酸基团脱去发生 β 消除反应后,加入合成的巯基季铵(2-dimethylaminoethanethiol),进行加成反应,然后在低能量的 CID 下,产生特异性的小分子碎片。前离子扫描技术大多数情况下,使用的是串联的质量分析器,如三级四极杆、Q-TOF 等,在后一级分析器上发现特征分子后,可以在同时的前一级分析器上即时的分析母离子,以便获取肽段的序列信息。在离子阱分析器上,较为常用的发现磷酸化肽段的方法则是作中性丢失扫描(neutral loss scan)。

(2)中性丢失扫描(neutral loss scan)

Hunter AP 等人最早阐述了磷酸化肽段在 CID 前后的质量数变化,在大多数情况下,磷酸化肽段的质量数变化是 98Da,以这一特征性的质量变化为判断磷酸化是否存在的依据[37]。中性丢失扫描在磷酸化质谱检测的方法中占有重要的地位,方法主要的优点在于思路比较简洁,可以完全在正离子模式下操作,不会像前离子扫描那样碰到负离子模式下的灵敏度下降的问题,此外,如前所说,此方法在离子阱分析器中也可以有很好的利用[38]。

中性丢失扫描也会遇到一些问题。因为质谱检测到的是质量电荷比而不是质量本身,因此在核对质量数的变化时,并不是只要核对 98Da 就可以。也就是说,中性丢失的信号可能有几个数值,在单电荷离子是 98Da,双电荷和三电荷是 49Da 和 32.66Da。这就提出一个要求,要么知道要检测的磷酸化肽段离子的电荷数,要么在软件中设定多个参数,前者对于未知物是难以了解的,而后者会增加程序分析的难度和准确性。更令人感兴趣的是,磷酸化肽段的电荷实际上与磷酸基团在 CID 中脱落的可能性有关。中性丢失另一个不可回避的问题是,磷酸化肽段质量数变化的多样性。在多数情况下,数值是 98Da,但也可能是 80Da,原因是磷酸化肽段在 CID 时发生的 β 消除反应时气态反应,反应中有可能失去的是一个磷酸分子(H_3PO_4),也可能失去的是 HPO_3,80Da[3]。不管是 98Da 还是 80Da,如果磷酸化的残基是酪氨酸,磷酸基团的脱落就不如另两种磷酸化残基容易。

从理论上说,在质谱的正离子模式下,使用中性丢失扫描是检测磷酸化肽段的好方法,在方法学的层面上,很早就有学者进行了研究,并认为是有效的[29,37]。但是,在面对未知的蛋白质样品时,就显得有些困难。原因是多方面的,最大的困难是在肽段混合物没有富集磷酸化的情况下,磷酸化肽段的信号与非磷酸化肽段的信号相比是较弱的,如果母离子的信号很弱,选择作 CID 时就会相当困难,这一点无论时手工选择或是程序选择都会有难度。当然,如果有富集的手段,或者离子化前的分离效果很好(比如使用多维色谱分离肽段混合物),还是会得到良好的效果的[20]。此外,中性丢失扫描的困难还在于,当磷酸化肽段太大时(>2 000Da),在作中性丢失扫描时,会发现信号的信噪比不好,这是因

为相对分子质量较大的肽段本身的信号就不好，更何况又有磷酸化。针对这一困难，有学者使用低特异性的蛋白酶（比如弹性蛋白酶，elastase）酶解磷酸化蛋白质，产生小分子的肽段以解决这一问题，肽段碎片序列的信息在一定程度上可以弥补低特异性蛋白酶造成的数据处理困难[38]。

以上介绍的是磷酸化肽段的检测技术，现在基本上是依赖于这两项技术对肽段的磷酸化与否进行判断。这两项技术各有优缺点。前离子扫描技术发展较早，有对负电荷的特征小分子碎片进行检测的，也有对正电荷的离子的，应用时选择余地相对较大，但是其负电荷模式对质谱灵敏度有较大的影响，而正电荷模式是应用于酪氨酸磷酸化，而且需要高分辨率的质谱（Q-TOF）。中性丢失扫描的技术特点是思路较为明确，可以应用于离子阱质谱，相信在磷酸化蛋白质组学中的应用会更广泛。

对蛋白质磷酸化的质谱分析分成两个方面，一是对寻找磷酸化肽段，二是对磷酸化的位置和数量进行分析。

（四）磷酸化位点的质谱分析

可以认为，磷酸化的位点分析是肽段磷酸化分析这个难点中的难点。

1. CID 碎片分析

一般的，蛋白质组学中要获得氨基酸序列信息，所依赖的是在质谱中获得相关肽段的碎片信息，然后使用一些算法（如著名的算法 SEQUEST 等）将碎片离子的质谱图谱与数据库进行比对、打分，获得统计结果，以确定肽段的序列。而这些碎片信息的获得，主要的是用 CID（collision induced dissociation），即所谓碰撞诱导分裂得到的。CID 技术是引入惰性气体将肽段离子撞碎，也就是说引入的是能量，是能级的升高迫使肽键断裂。磷酸基团的酯键的键能比肽键的要低，因此，在 CID 时，磷酸酯键比肽键更容易断裂，这一点直接造成了磷酸化位点判断的困难。

尽管磷酸酯键在 CID 时是相对脆弱的，但是近年来仍然有很多小组致力于使用这项技术结合中性丢失扫描技术等，研究蛋白质的磷酸化状态[32,38]。也有学者报道，在肽段氨基末端进行化学修饰，会使较多的磷酸基团保留在肽段碎片中，使磷酸化肽段的位点分析更为简单[39]。

也有使用 MALDI 离子源，配上新颖的质量分析器，如 Qq-TOF、TOF-TOF 和传统的离子阱质谱进行 CID 碎片分析。有学者报道，在 MALDI-ION TRAP 中，酪氨酸磷酸化肽段的磷酸基团在 CID 中相对不容易脱落，而且脱落与肽段中碱性氨基酸的存在有关[26]。其实磷酸化肽段的磷酸基团在 CID 中也有可能不发生脱落，只是这种保留是不常见的，而且与 CID 的能量大小很有关系，实际工作中 CID 能量基本是不变的[40]。在 CID 分析中最大的问题就是磷酸基团的脱落，这种脱落有时造成磷酸化位点的判断几乎是不可能的。因此，通过化学修饰，使磷酸基团转化成相对稳定的基团，在进行低能量的 CID 时尽可能保留磷酸化位点（尽管此时已经不是磷酸基团了），可能对磷酸化的位点信息进行了解。已有学者对此展开了研究，并获得了进展[41]，最近已有文献报道使用对 β 消除反应后的磷酸化肽段进行亚硫酸盐加成，使得磷酸基团转成亚硫酸根，提高了这个位点上基团的稳定性[42]。图 18.5 是这种化学方法的示意图。这个工作小组的试验解决了 CID 中磷

酸基团脱落带来的磷酸化位点分析的困难。这种方法是对丝氨酸和苏氨酸磷酸化有效，因为酪氨酸磷酸化肽段不能发生β消除反应，所以对它是无效的。另外，亚硫酸基团的引入，也造成了CID图谱分析难度的加大。尽管如此，这种以及今后可能出现的其他种类的化学修饰法，在一定程度上稳定了磷酸化位点，是蛋白质组学中立足于现有质谱解决磷酸化的位点明确的有前景的策略。

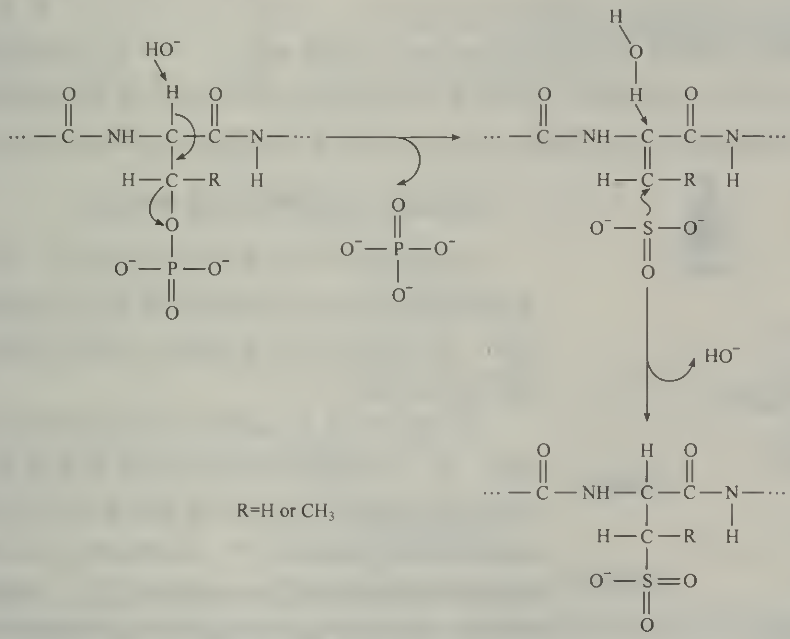


图 18.5　对丝氨酸和苏氨酸磷酸化的亚硫酸盐加成化学过程

2. ECD(electron capture dissociation，电子俘获分裂)**碎片分析**

FTICR，即傅里叶离子回旋共振质谱，是一种相对较新的用于生物合成的质谱，具有相当高的分辨率和准确性，是几乎现有的生物质谱中功能最强大的一种。对于磷酸化肽段的分析，FTICR也有非常重要的作用，其原因就是FTICR的碎片模式是所谓电子俘获分裂的技术，其原理是照射一束亚热态(subthermal)的电子到电喷雾所产生的磷酸化肽段，或者是小分子的蛋白质本身，使其形成碎片[43]。与CID技术撞击肽段时产生的大量b和y族碎片不同，ECD技术产生的多是c和z族碎片，更为重要的是，磷酸基团在ECD中保留在碎片分子中，使得直接判断磷酸化位点成为一种可能[8]。FTICR的分辨率是极高的，一般是其他生物质谱的10倍左右。正因为这一点，它甚至可以对小分子蛋白质不进行酶解而直接分析磷酸化状态[44]。随着各大质谱公司相继推出各自的FTICR，相信ECD技术不破坏磷酸化位点的特点一定会被更多地应用到实际工作中。

(五) 磷酸化蛋白质组学中的定量问题

不管是蛋白质组学的哪一个分支，都会碰到定量的问题，蛋白质磷酸化的定量更是其中的难点。蛋白质组学的定量问题，一直存在疑问。因为生物体中的蛋白质作为一个整

体而存在，大多数蛋白质在生命活动中实际上是动态的。这样，定量问题从根本上说就不是一个数值能够代表的，应该是一个动力学过程。蛋白质磷酸化更是带有这个特征，生命周期中的蛋白质磷酸化是具有重要作用的一种修饰，也是非常复杂的，大多数相当重要的蛋白质磷酸化是"短命"的，要想分析这些也许变化迅速的磷酸化状态，从理论上而言是困难的。

从质谱的技术角度而言，磷酸化的定量仍然是个挑战。已经有许多工作致力于使用质谱技术解决定量问题，其中有的尝试是使用液质联用技术[45,46]，先分离磷酸化肽段然后比较其在质谱上的信号强度，以求定量估计磷酸化的相对程度。但是，序列对磷酸化和非磷酸化肽段的质谱信号的影响是复杂的，所知寥寥，这些努力的有效性值得怀疑。

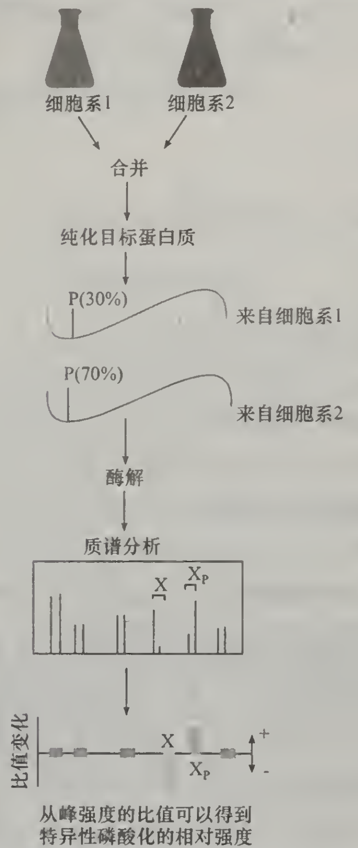


图 18.6 同位素标记的培养基培养的细胞系用于磷酸化蛋白质的定量分析示意图

磷酸化肽段 X 被示意为从细胞培养基 1 的 30%磷酸化，变化到细胞培养基 2 的 70%磷酸化。在质谱图上的反映是非磷酸化肽段 X 的同位素相对比例降低(后者比前者)，而同时测到的磷酸化肽段 X_P 比例升高

同位素标记在磷酸化定量中的应用

生物学研究中同位素的应用历史悠久，在原子物理发展的早期，生物学家就将同位素应用于生物学的各个领域。蛋白质组学的定量问题的主要技术思路，仍然与同位素标记有关。

从目前研究来看，磷酸化研究的同位素标记有两种类型。第一种研究策略是在细胞系的培养基中加入稳定同位素(主要是^{15}N)，在普通培养基中生长的细胞进行比较而得到定量的结果[47]。这种试验方法具体操作是将两种不同的培养基下的细胞等量混合，裂解细胞后分离所感兴趣的蛋白质，经酶解，并由质谱检测分析。选定的蛋白质的非磷酸化肽段的同位素峰比例是相对固定的(并不是一定相等，而是比例大致不变)，而同时磷酸化肽段及其非磷酸化部分的比例关系会发生变化，这种变化应该是向相反方向发展的，详见图 18.6。这种方法的设计可谓巧妙，让人看到了同位素标记技术和质谱技术的结合产生的"化学反应"式的效果，这种方法将蛋白质的鉴定和磷酸化的定量同时完成，而且还可以了解磷酸化的程度，一举多得。但是，这种思路的缺点也是明显的。首先，^{15}N 同位素标记的培养基的获得有一定的困难；其次，这种方法只能应用于培养细胞的研究，对其他生物样本有困难。标记的效率和均一性也是一个问题。另外，质谱分析前需要对特定蛋白质进行分离，低丰度的蛋白质能否实现有效的分离，值得怀疑。分离酶解后是否能找到感兴趣的磷酸化肽段也并非完全由试验者的意愿决定。尽管如此，这种方法仍然是较早在磷酸化定量领域中应用同位素技术的尝试，为其他此类工作的开展开拓了思路。

第二种类型的同位素策略是使用同位素标记的亲

和试剂和磷酸化肽段的磷酸化位点结合，从而给磷酸化肽段加上亲和尾部，引入质量数的差异[48,49]，以便在质谱分析时根据信号的强度进行定量。这种方法的化学过程示意图见图 18.3。该方法的思路实际上来源于 ICAT，因此被称为 PhIAT(phosphoprotein isotope-coded affinity tag)，只是肽段上的磷酸基团参加反应而不是巯基。这种定量思路比前者要直接一些，更重要的是实验所用的 PhIAT 试剂，要想合成对于生物学实验室是有难度的，而购买商品化的试剂比较昂贵。但是，从 ICAT 方法在蛋白质组学的一般定量中的应用来看，PhIAT 应该有良好的前景。PhIAT 的示意图见图 18.7。

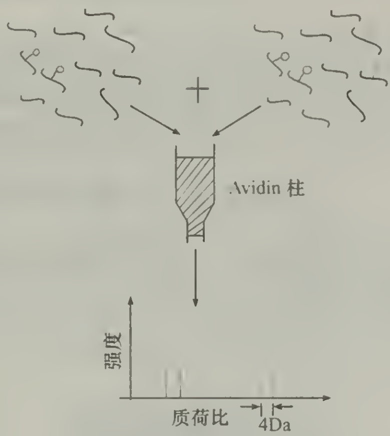


图 18.7　PhIAT 的流程示意图

图中弯线表示肽段，带棒小球表示同位素亲和尾部。下部质谱图谱示意磷酸化肽段的相对数量

(六) 蛋白质磷酸化的动力学研究

蛋白质的磷酸化过程是一个动态的过程，因此可以说静态的磷酸化定量研究是不完善的，磷酸化的研究在技术条件允许的情况下，总是要进行动态、定量研究的。现在，对蛋白质磷酸化的动力学研究才刚刚开始，技术上还处于初级阶段，不过已经有学者进行了探索。Ruse 等人使用的是液质联用，计算的是色谱峰的面积，在有标准肽段进行校正下，对体外的磷酸化反应进行了动力学分析[50]，不过他们的工作只是限于标准品。总的来说，磷酸化的动力学研究还有待后来者开拓。

二、材　　料

(一) 试剂和溶剂

①所有缓冲液均用经过 Millipore (Bedford, MA, USA)系统处理过的水配制。

②冰醋酸，购自 Sigma (St. Louis, MO, USA)公司。

③碳酸氢铵，购自 Sigma (St. Louis, MO, USA)公司。

④乙腈 (ACN)，购自 Fisher (Fair Lawn, NJ, USA)公司。

⑤三氟乙酸 (TFA)，购自 Merck (Schuchardt, Germany)公司。

⑥三氯化铁($FeCl_3$),购自 Sigma (St. Louis, MO, USA)公司。
⑦氯化钠(NaCl),购自 Sigma (St. Louis, MO, USA)公司。
⑧α-腈基-4-羟基肉桂酸(HCCA),购自 Sigma (St. Louis, MO, USA)公司。
⑨Poros 20 MC 填料,购自 ABI (St. Louis, MO, USA)公司。
⑩各种内径的石英毛细管,购自河北省永年县色谱设备有限公司。
⑪甲醇,购自 Fisher (Fair Lawn, NJ, USA)公司。

(二)蛋白质和酶

①胰蛋白质酶,购自 Roche (Mannheim, Germany)公司。
②牛 β-酪蛋白 (β-casein),购自 Sigma (St. Louis, MO, USA)公司。
③β-lactoglubulin,购自 Sigma (St. Louis, MO, USA)公司。

(三)仪　　器

试验所用的质谱仪为 Q-TOF Ultima-global™(Micromass,UK)。

三、方　　法

(一)胰蛋白酶酶解

100μg 蛋白质溶于 30μl NH_4HCO_3(100mmol/L, pH 7.8)缓冲液中。先用 DTT (终浓度为 5mmol/L)在 50℃ 条件下还原 15min,冷却至室温后,加入 IAA (终浓度为 10mmol/L)在室温下再反应 15min (避光)。用 NH_4HCO_3缓冲液稀释到 100μl,加入浓度 1μg/μl的胰蛋白酶溶液 2.5μl,37℃保温 4h。补加酶 2.5μl 后,37℃继续保温 18h。偶尔取出轻摇。酶解完成后取出,置于 −20℃ 保存。

(二)标准蛋白质酶解产物中磷酸化肽段的 IMAC 纯化

本实验室进行了 IMAC 方法在标准蛋白质酶解产物中的应用。IMAC 方法的关键之处在于金属螯和柱的装填,本实验室使用自己装填的金属螯和柱应用于磷酸化研究中,并取得了初步的成功,目前正在进行更深入的应用试验,期待能在实际工作中发挥其作用。

1.IMAC 柱的装填和活化

使用 POROS 20MC 填料,将适量干填料混悬于水中(约 1mg/ml),振荡数分钟后,离心去除上清,加入 0.1%乙酸(约 1mg/ml)。取一段长度在 10cm 左右、内径 320μm 的石英毛细管,依次用 10 倍柱体积的水、0.1mol/L NaOH、0.1mol/L HCl 和水冲洗石英毛细管内壁。在毛细管一端塞入少许填料,火焰瞬间烧灼,形成烧结物,显微镜下观察烧结物的形成情况,确信烧结物已经形成,然后以 10 倍柱体积的水冲洗毛细管。将填料混悬液振荡均匀,灌入毛细管中,直到装填柱的长度达到约 6cm。

依次用 10μl 50mmol/L EDTA、40μl 0.1mol/L 乙酸冲洗 IMAC 柱,用 100μl 0.1mol/L $FeCl_3$ 活化 IMAC 柱,然后用 20μl 0.1mol/L 乙酸冲洗。

2. 对磷酸化肽段的纯化

样品(<100μg)溶解于 10μl 0.1mol/L 乙酸中,缓慢上柱(1μl /min)。依次用 20μl 0.1mol/L 乙酸、25μl 25% 乙腈(1%乙酸)和 20μl 0.1mol/L 乙酸冲洗,去除非特异性结合的肽段,然后用 10μl pH 10.5 氨水洗脱两次,再合并洗脱液,加 20μl 0.1mol/L 乙酸酸化后冻干。

3. 磷酸化肽段的质谱分析

对磷酸化肽段进行质谱分析前,应该脱盐。对于蛋白质量少的样品(<5μg),使用 ZipTip™(C_{18}, Millipore)脱盐。如果蛋白质量大于 5μg,则使用自己装填的 C_{18}柱脱盐(方法和上文 IMAC 类似)。脱盐后进行质谱分析。

4. 甲酯化反应在磷酸化肽段分析中的应用

冻干酶解后样品。160μl 乙酰氯缓慢加入 1ml 甲醇中,生成 meth.HCl。5min 内,加若干微升 2mol/L meth.HCl 到酶解产物中(以 1μg 酶解产物加 1μl 的比例加入),室温反应 2h。冻干反应物,重复上述过程 2 遍,冻干反应物,加 10μl 0.1mol/L 乙酸溶解,继续进行 IMAC 纯化。

四、结　　果

(一) IMAC 对 β-酪蛋白酶解产物中的磷酸化肽段的纯化效果

本实验室试验了自制的 IMAC 柱对标准蛋白质 β-酪蛋白酶解肽段中的磷酸化肽段的纯化效果,进行了 IMAC 纯化物的 MALDI-QTOF 分析。并对甲酯化反应的条件进行了初步的试验。

1. MALDI-QTOF 分析结果

(1) IMAC 纯化后的质谱图

IMAC 纯化后的图谱大大简化,仅剩下磷酸化肽段(m/z 2061.2)和其脱磷酸的分子(m/z 1963.3)。如图 18.8 所示。

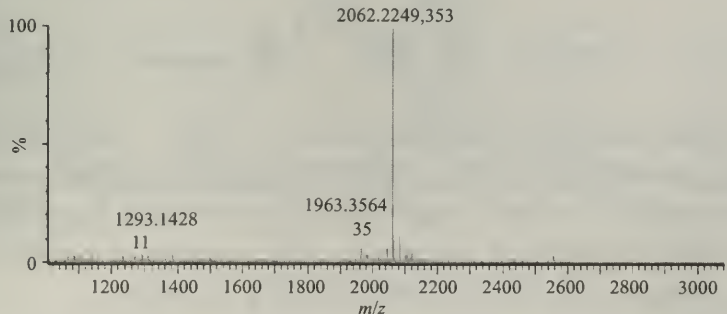


图 18.8　标准蛋白质 β-酪蛋白酶解肽段 IMAC 纯化后的质谱图

(2) 磷酸化肽段(m/z 2062.2)的 MS/MS 碎片图谱

另外还进行了磷酸化肽段(m/z 2062.2)的二级碎片质谱,如图 18.9 所示。

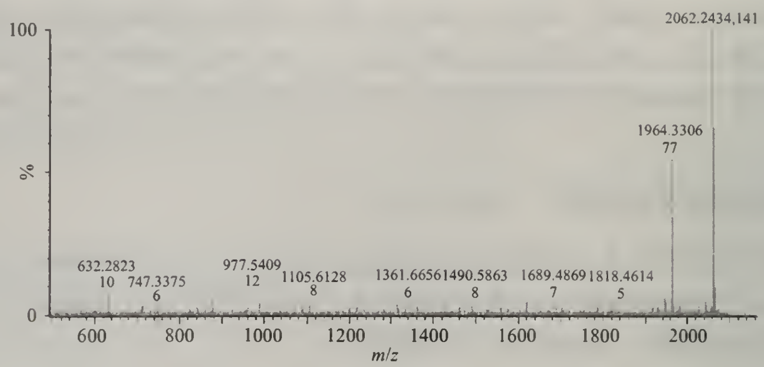


图 18.9 磷酸化肽段(m/z 2062.2)的 MS/MS 碎片图谱

2. 酯化反应后的 IMAC 纯化肽段在 MALDI-QTOF 上的分析结果

甲酯化反应显示了肽段上的羧基位点被逐个甲酯化的过程,这里显示的是仅仅采用一步反应的图谱。实际工作中应重复甲酯化反应,以提高甲酯化的效率,尽量做到完全反应。见图 18.10。

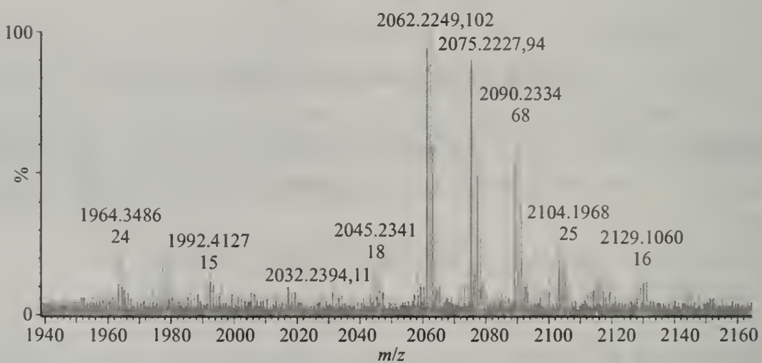


图 18.10 酯化反应后的 IMAC 纯化肽段在 MALDI-QTOF 上的分析结果

五、展　　望

蛋白质的磷酸化研究是蛋白质组学诞生以来,第一次将一个整体的研究对象对准一项翻译后修饰,选择磷酸化是有道理的。磷酸化是生命活动中最重要的一项翻译后修饰,与信号传导、细胞周期、生长发育以及癌症机制等诸多生物学问题有着密切的关系。但同时,磷酸化研究也可以说是蛋白质组学遇到的一个真正的难题,而且是无法回避的,不能充分理解蛋白质磷酸化现象,就无法认识蛋白质在生命中的整体意义。

就现阶段磷酸化研究的情况来看,磷酸化研究还远远没有达到常规试验的水平,现在

所做的仍然是为不久的将来大规模开展磷酸化蛋白质组学研究做技术准备。

大多数情况下，磷酸化蛋白质的研究就是磷酸化肽段的研究。磷酸化研究的特点如下：第一，磷酸化蛋白质的丰度比非磷酸化蛋白质要低，相对的磷酸化肽段的丰度也要低于非磷酸化肽段；第二，磷酸化蛋白质的磷酸基团在样品的制备过程中有可能失去；第三，磷酸化蛋白质的胰蛋白酶酶解的效率比非磷酸化的蛋白质要低，这一点主要是因为磷酸化的负电性会干扰胰蛋白酶的作用[3]；第四，在质谱信号中，磷酸化肽段的负电性造成离子化困难，而且非磷酸化肽段的信号还会抑制磷酸化肽段的信号。因此，磷酸化蛋白质研究中的样品制备非常重要，除了在制备时加磷酸酯酶抑制剂等措施外，磷酸化蛋白和肽段的富集是很必要和关键的。现在富集的方法包括磷酸化特异性抗体和 IMAC，其中磷酸化肽段甲基酯化后的 IMAC 方法最为有效[20]。另外，各种各样的化学修饰在磷酸化蛋白质组中的应用是很有效果的。因此，要研究磷酸化，选择合理有效的化学策略是必要的[42]。在不久的将来，磷酸化蛋白质组的技术积累达到一定程度后，就会使用这些方法来解决具体的生物学问题，到那时，磷酸化蛋白质组将真正成为一项大事业。磷酸化蛋白质组研究的策略汇总见图 18.11。

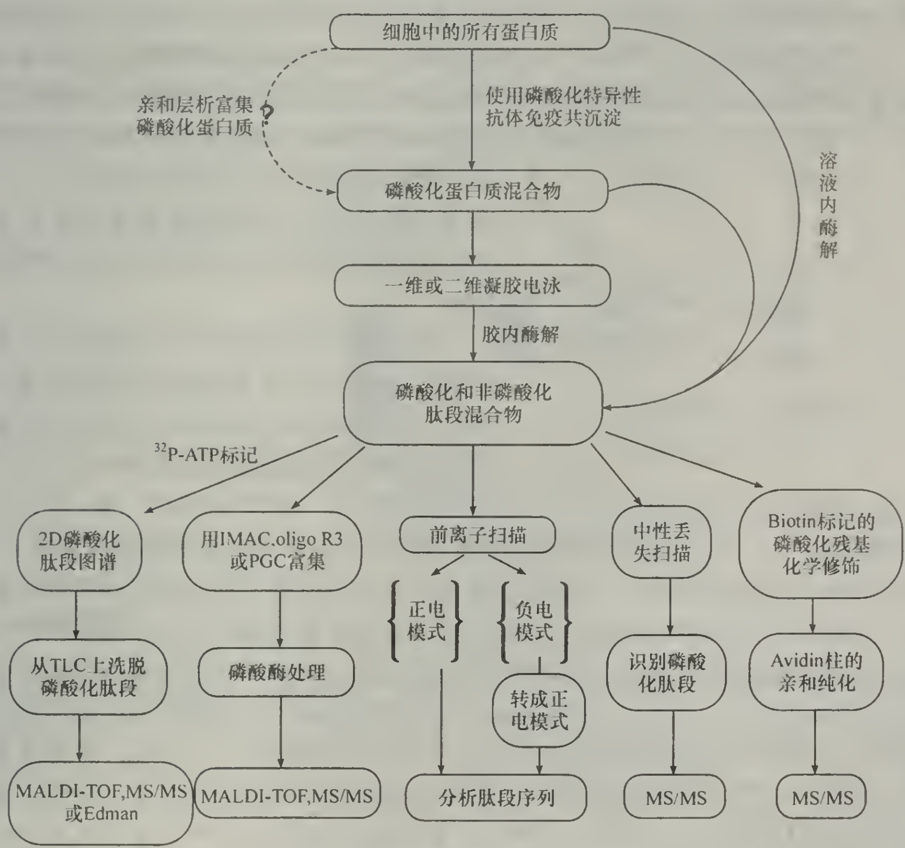


图 18.11　磷酸化蛋白质组研究的策略汇总

第二节　糖基化蛋白质组研究

一、介　　绍

（一）概　　况

1988年，牛津大学的生物化学家Raymond Dwek在《生化年评》(*Annual Review of Biochemistry*)上，发表了题为"糖生物学"的综述，标志着糖生物学这一新的研究领域的诞生[51]。近几年来，随着不同物种基因组计划的陆续完成，以及蛋白质组研究的开展和各种分离鉴定技术的进步，推动了糖生物学的迅速发展。美国新罕布什大学(UNH)化学系的结构生物学研究中心启动了一项"线虫糖组学研究"计划，分析测定线虫的糖组(glycome，定义为细胞内所有的糖链组分)，开展糖组学(glycomics)研究，目的在于确定基因所携带的遗传信息与其功能之间的关系。人类基因组计划完成后，美国于2001年9月正式启动了"功能糖组学"研究项目，项目的总体目标是阐明由蛋白质-糖链相互作用所介导的细胞通讯机制[52]。基因组计划提供了最基本的遗传信息，而许多基因的功能仍需要阐明。其中的关键就是要了解蛋白质的翻译后修饰，蛋白质糖基化便是最主要的翻译后修饰之一。尽管糖组的定义与基因组和蛋白质组的定义类似，但其研究应该主要集中在蛋白质糖基化研究上。而蛋白质糖基化研究的基本目标是：寻找编码糖蛋白的基因(即基因组信息)；寻找糖基化位点(即糖基化位点信息)；解析糖链及糖基化肽段的结构(即结构信息)；研究糖基化的功能(即功能信息)[53]。为了实现这些目标，应当更好地将蛋白质组学的技术和研究策略应用到大规模的蛋白质糖基化研究中。

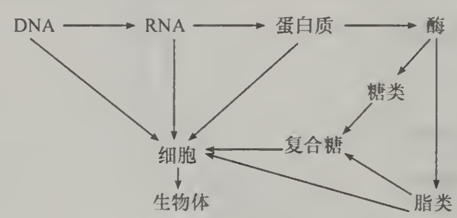


图18.12　生物信息流

现代分子生物学的中心法则认为，生物信息的流向是从DNA到RNA再到蛋白质。而实际上，糖类和脂类也是组成生物体所必需的两类主要分子，生物信息流见图18.12所示[54]。

因此，我们应当从基因组水平来关注蛋白质糖基化。在真核细胞中，有一半以上的蛋白质被糖基化[55]。首先，所有生物的细胞表面都由很多不同类型的糖链所包被，并且细胞内也存在各种类型的糖基化；蛋白质糖基化在蛋白质折叠、蛋白质定位和转运、蛋白质溶解性、抗原性以及细胞与细胞的相互作用等方面都有着重要的作用。最近发现构成细胞膜的许多脂肪或蛋白质是以糖脂或糖蛋白形式存在，而且糖低聚物在细胞间信号传递中也起了重要的作用。另外，由于糖在HIV或其他致病源与血细胞结合过程中扮演了重要的角色，它成为免疫系统控制和癌症治疗的最重要的线索之一。其次，具有足够的多样性是作为任何信息分子的基础，糖链比核酸和蛋白质在结构上更具有可变性。聚糖极高的复杂性明显地归因于"连接方式异构体"和"分枝形成"，而这种现象在其他的生物信息分子中并不存在。例如，由6个核苷酸组成的寡核苷酸，可能的序列仅有4 096种(4^6)；6个氨基酸组成的六肽，可能的序列有6.4×10^7种(20^6)；而由6个单糖组成的寡糖链，可能的序列则多达1.05×10^{12}种[56]。由于糖链结构的复杂性，其可能包含的信息量

比核酸和蛋白质增加了几个数量级。第三，聚糖中许多令人感兴趣的方面可能反映了它们的生物和分子的进化作用。例如，多糖中少数稳定构成单位的自然选择；甲醛聚糖(formose)反应可能是前生物期合成糖类的反应，而糖酵解逆反应中醇醛缩合并生成果糖的过程又和甲醛聚糖反应的最后一步相同；Lobry-de-Bruyn 重排反映了果糖、葡萄糖及甘露糖之间的亲缘关系；复杂的单糖如唾液酸和脱氧己糖可由葡萄糖或甘露糖产生，反映了生物合成的保守性；乳糖的出现可能是在分子进化过程的后期；核糖和核酸的起源仍是一个未解决的问题，等等[57,58]。

在真核生物细胞中，蛋白质糖基化分类是基于糖链和蛋白质的连接类型的。可分为以下三类。

1．*N*-糖基化

通常是指糖链共价连接在蛋白质的天冬酰胺(Asn)的氨基侧链处。糖链与天冬酰胺(Asn)残基的连接发生在三个残基的特殊识别序列(Asn-X-Ser/Thr，其中 X 为除 Pro 和 Asp 以外的任何氨基酸)处。*N*-糖链通常被划分为三种：高甘露糖型、复杂型和混合型。所有的 *N*-糖链都有一个相同的五糖核心 $Man_{\alpha1\text{-}3}(Man_{\alpha1\text{-}6})Man_{\beta1\text{-}4}GlcNAc_{\beta1\text{-}4}GlcNAc$ (GlcNAc2Man3)(图 18.13)。其中高甘露糖型可在这个五糖核心上再连接 2～9 个甘露糖(mannose)；复杂糖型则有更多的 *N*-乙酰葡糖胺(*N*-GlcNAc)、半乳糖(galactose)、唾液酸(sialic acid)、岩藻糖(fucose)及木糖(xylase)与五糖核心相连，其中唾液酸提供了额外的负电荷；混合型的糖链结构则兼具以上两种糖链结构的特点[53]。

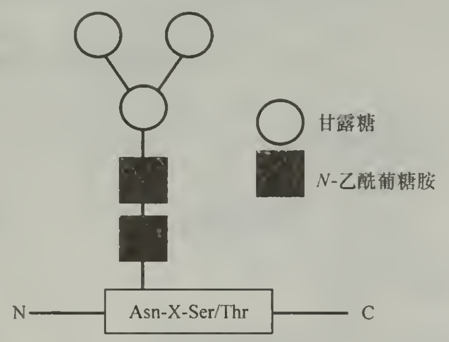


图 18.13　*N*-连接糖链的五糖核心结构

2．*O*-糖基化

通常是指糖链通过 *N*-乙酰葡糖胺连接在蛋白质的丝氨酸或苏氨酸侧链的羟基处，也有其他类型的连接，如通过 *O*-甘露糖。其中，多细胞真核生物中存在细胞核和细胞质蛋白质丝氨酸或苏氨酸残基的动态糖基化(*O* 连接的 *N*-乙酰葡糖胺即 *O*-GlcNAc)。在一些蛋白质中，*O*-GlcNAc 糖基化与 *O*-磷酸化发生在相同或相近的位点上，从而出现一种假说，认为 *O*-GlcNAc 可暂时阻断 *O*-磷酸化。蛋白质的 *O*-GlcNAc 在生理、病理以及信号转导等过程中都有重要作用[59]。

3. 糖基磷脂酰肌醇锚(GPI-Anchor)糖基化

通常是指蛋白质通过C端共价连接的糖基磷脂酰肌醇锚定在膜脂上，这将细胞外蛋白质和细胞膜相连。通过GPI连接到细胞膜上的蛋白质与穿越细胞膜的蛋白质不同，因为它们可被PI-PLC(磷酸肽肌糖特异性磷酯酶C)分离到细胞外的空间。在真核细胞中，很多胞外蛋白质(包括一些水解酶、细胞表面抗原、细胞黏着蛋白等)都通过糖基磷脂酰肌醇锚结合在质膜上[60]。

(二) 研究蛋白质糖基化的常用方法

与核酸和蛋白质不同，糖链很少以线性、非分支的方式出现，而是存在许多可变的修饰。因此，对糖链完整序列的测定很难通过一种方法来完成，而是需要物理和化学方法的反复组合，最终确定糖链详细的结构信息(如质谱和核磁共振的方法)，图18.14展示了糖蛋白分析的流程[61]。研究糖基化的其他工具和方法包括酶(内切糖苷酶和外切糖苷酶)、凝集素(糖结合蛋白)、化学的修饰及断裂、新陈代谢放射性标记、糖基化抑制剂、抗体、糖基化转移酶的分子克隆以及在生物体及细胞中的遗传操作等[54]。

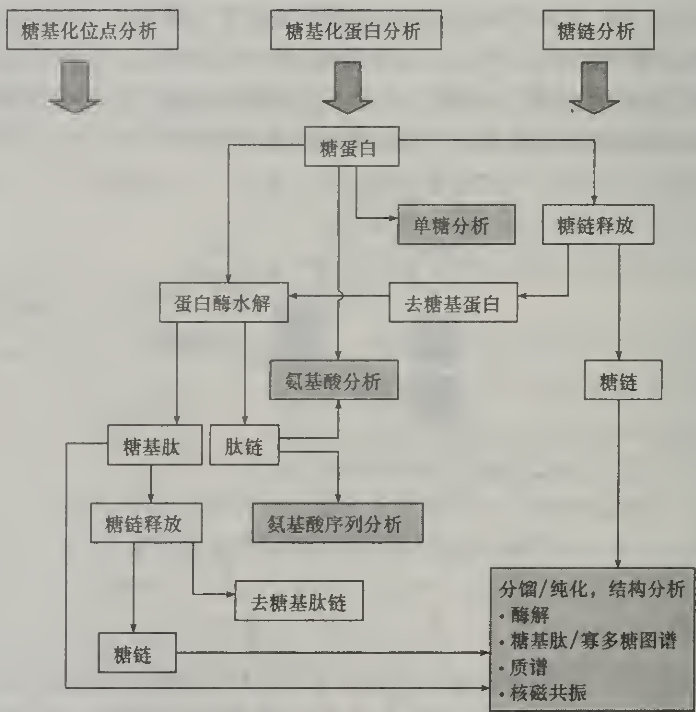


图18.14 糖蛋白分析的流程图

(三) 糖基化蛋白质的分离技术

从“基因组学、蛋白质组学和糖组学”的整体观念出发，应以基因组信息为基础，以研究糖蛋白为中心。糖基化蛋白质组学研究糖肽比研究单纯的聚糖更合适，其首要的任务

是从蛋白质混合物中分离纯化糖蛋白，再从糖蛋白中分离纯化糖肽，可以采用“糖捕获”法[57,58]得到。多种凝集素可被用来分离带有不同糖链结构的糖蛋白和糖肽，如半乳糖苷结合凝集素(galectin)可分离包含 *N*-乙酰乳糖胺(LacNAc)的 *N*-连接的和 *O*-连接的糖链；伴刀豆凝集素 A(ConA)可分离 *N*-连接的糖链。花生凝集素(PNA)可分离包含 T 抗原(Galâ1-3GalNAc 结构)的 *O*-连接的糖链；*Aleuria aurantia* 凝集素(AAL)对含有 L-岩藻糖的糖链具有广泛的特异性，等等。如何选择这些凝集素在“糖捕获”法中是最为关键的步骤。正因为如此，有必要了解一些具有代表性的凝集素的典型特征，如亲和性、特异性、热稳定性、是否需要金属离子和多共价键的情况等。伴刀豆凝集素 A 和 galectin LEC-6(Gal6)特异性地结合到高甘露糖型或复杂型的 *N*-连接的糖链；而花生凝集素(PNA)则对存在于 *O*-连接的糖链的 T 抗原(Galβ1-3GalNAc 结构)具有很高的亲和性。

然而，应当注意到伴刀豆凝集素 A 即其他类型的多价植物凝集素对一系列的糖链都具有很高的亲和性，如亲和常数 K_a 在 $10^6 \sim 10^8$L/mol 之间。但 galectin 对包含乳糖胺的糖链的亲和常数 K_a 却很低，在 $10^3 \sim 10^6$L/mol 之间。因此，那些亲和性太低的糖蛋白(或糖肽)有可能没有被一些 galectin 柱吸附。伴刀豆凝集素 A、galectin LEC-6(Gal6)、花生凝集素是常用的三类凝集素，在此分别阐述其主要用途及功能[53]。

1. 伴刀豆凝集素 A(ConA)

伴刀豆凝集素 A 是一种典型的对于一系列高甘露糖型的 *N*-连接的糖链有高亲和性的植物凝集素。伴刀豆凝集素 A 同时也能结合复杂型的 *N*-连接的糖链及双分枝的复杂型的 *N*-连接的糖链。通过一系列的糖链分析表明，与伴刀豆凝集素 A 结合的高甘露糖型糖链结构可分为三组(如图 18.15)：伴刀豆凝集素 A 对于含有三甘露糖核心结构“Manα1-6(Manα1-3)Manβ”的糖链具有较强的亲和性，亲和常数 *K*a 范围为 $1\times10^6 \sim 4\times10^6$ L/mol，这组包括 M3B、M4B、M5A、M6B、M6C、M7B、M7D 和 M8C；末端非还原性甘露糖以 α1-2 连接到 Man(α1-6)上，从而使得三甘露糖核心结构“Manα1-6(Manα1-3)Manβ” 与伴刀豆凝集素 A 的亲和性增强了一个数量级。伴刀豆凝集素 A 对于含有四甘露糖核心结构“Manα1-2Manα1-6(Manα1-3)Manβ”的糖链亲和力最强，亲和常数 *K*a 范围为

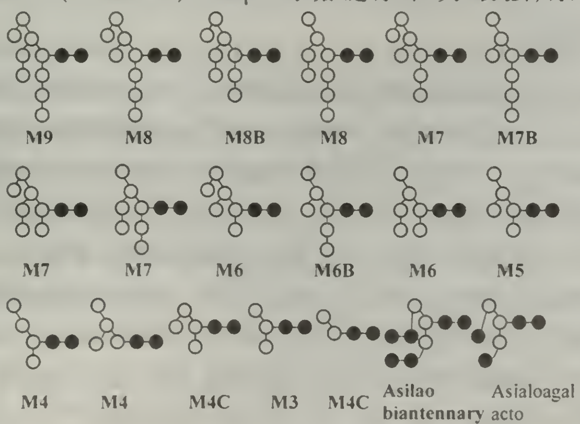


图 18.15 与伴刀豆凝集素 A 结合的三组高甘露糖型糖链结构

$3\times10^7\sim5\times10^7mol/L^{-1}$,这组结构包括 M6A、M7A、M7C、M8A、M8B 和 M9A;而缺少三甘露糖核心结构,或含有阻碍性的 3-OH 及 6-OH 取代基团的 Man(α1-6)的糖链则对伴刀豆凝集素 A 具有较弱的亲和性,这类糖链为第三组。

2. galectin LEC-6(Gal6)

galectin 是一类分布广泛、不依赖于金属离子的可溶性凝集素,其对 β-半乳糖苷的特异性在很多多细胞生物中具有保守性。到目前为止,已有 12 种哺乳动物的 galectin 被发现,并被命名为 galectin-1 到 galectin-6。根据它们的结构,可分为三类:原型、嵌和体型和串联重复型。在线虫(*C. elgans*)中,galectin 被命名为 LEC1-11,galectin LEC-6(Gal6)代表了线虫中主要的一类 galectin,其生化性质得到了广泛的研究。通过前沿亲和层析法(FAC)[62]对 Gal6 的结合特异性进行了研究,结果表明:Gal6 和具有分枝结构的复杂型 *N*-连接的糖链结合力最强。随着分枝的增加,亲和常数 K_a 也随之增加,这是由于糖链的化合价增加的结果,而末端非还原性的 α2-6 唾液酸化消除了这种亲和性[63]。Gal6 最倾向于结合包含半乳糖且具有分枝结构的 *N*-连接的糖链。

3. 花生凝集素(PNA)

花生凝集素在组织化学中是用来检测 *T* 抗原(Galβ1-3GalNAc 结构)的,T 抗原在黏蛋白(如白细胞唾液酸蛋白、血型糖蛋白 A)的 *O*-连接的糖链中广泛存在。花生凝集素一般选择性地结合到广泛存在于 *O*-连接的糖链的 *T* 抗原上。

(四)糖基化蛋白质的检测技术

一般地,可以将^3H 或^{14}C 标记的糖加入培养的细胞或组织中,通过放射性自显影检测糖蛋白。此种方法灵敏度较高,但放射性标记的过程缓慢、危险,并且花费较大。检测糖蛋白的非放射性方法常采用高碘酸-希夫碱(P.A.S)反应,通过颜色变化来检测糖蛋白,还有采用阿利新蓝(Alcian blue)染色等。但这些方法比考染的灵敏度还低,很难用于当前蛋白质组学研究中[64]。下面介绍几种常用的检测方法。

1. 荧光染色检测方法

丹磺酰肼作为酸性品红染料的荧光替代试剂可用来检测聚丙烯酰胺凝胶中的糖蛋白。胶中的糖蛋白先经过高碘酸氧化,形成醛,然后与丹磺酰肼反应,形成腙,接着用硼氢化钠还原成肼的衍生物。这一连串的反应需要将胶置于酸化的二甲亚砜中 60℃温育 2h,并且需要高浓度的丹磺酰肼染料(0.5~2.0g/L),所以此方法需改进[65]。电印迹到聚偏二氟乙烯(PVDF)膜上的糖蛋白可通过荧光素氨基脲酰肼和 8-萘胺-1,3,6-三磺酸盐来检测,但不适于直接检测聚丙烯酰胺凝胶中的糖蛋白[64]。荧光素氨基脲标记的方法没有得到广泛应用,而 8-萘胺-1,3,6-三磺酸盐需要在 60℃温育 45~60min,灵敏度一般为 50~200ng[66]。还有一种灵敏的绿色荧光染色糖蛋白检测方法,即用 Pro-Q Emerald 300 染色检测聚丙烯酰胺凝胶中的糖蛋白或聚偏二氟乙烯膜上的糖蛋白[67]。此过程需固定、高碘酸氧化、Pro-Q Emerald 试剂染色三个步骤(图 18.16)。Pro-Q Emerald 300 试剂和 Pro-Q Emerald 488 试剂对胶中糖蛋白的检测灵敏度分别可达 1ng/带和 4ng/带,比标准高碘酸-

希夫碱(P.A.S)方法灵敏约 50 倍;而对印迹在聚偏二氟乙烯膜上的糖蛋白检测灵敏度则为每条带 2~18ng。Pro-Q Emerald 300 试剂染色的胶还可用 SYPRO Ruby 试剂复染,SYPRO Ruby 试剂对总蛋白质染色呈橙色,Pro-Q Emerald 300 试剂对糖蛋白染色呈绿色。

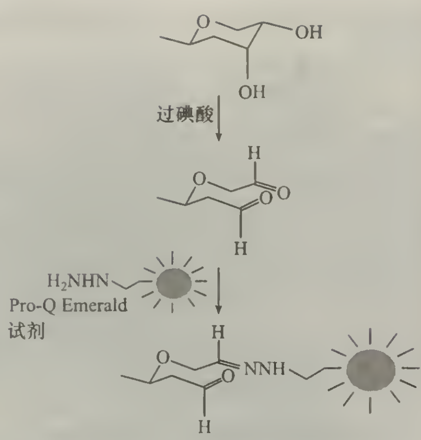


图 18.16 Pro-Q Emerald 试剂染色检测糖蛋白的三个步骤

2. 与凝集素联用的染色检测方法

蛋白质经过聚丙烯酰胺凝胶分离,印迹到聚偏二氟乙烯膜上,然后与结合了碱性磷酸酯酶的凝集素一起孵育。再通过碱性磷酸酯酶酶解 DDAO 磷酸盐,生成长波的红色荧光的 DDAO 来检测。此方法检测糖蛋白的灵敏度约为 15ng,依赖于蛋白质的性质和糖基化程度。Pro-Q 有三类凝集素试剂盒:伴刀豆凝集素 A 用于检测末端有 α-甘露糖吡喃糖苷残基或 α-葡萄糖吡喃糖苷残基且不具有三天线或四天线结构的糖蛋白;小麦胚芽凝集素(WGA)用于检测有 *N*-乙酰葡糖胺或 *N*-乙酰神经氨酸的糖蛋白;西非单叶豆凝集素(*Griffonia simplicifolia* lectin Ⅱ,GS-Ⅱ)用于检测末端有非还原性的 *N*-乙酰葡糖胺的糖蛋白。还可以用结合了辣根过氧化物酶(HRP)的伴刀豆凝集素 A 或小麦胚芽凝集素等标记糖蛋白,再用显色剂(DAB,H_2O_2)显色[68,69]。

3. 其他检测方法

可用酶解或化学裂解方法将糖链从糖蛋白中释放出来,经过 2-氨基苯甲酸等标记,用毛细管电泳或高效液相色谱等分离,再由检测器检测[70],以及用质谱、核磁共振(NMR)等方法来进行检测。

(五) 质谱法在蛋白质糖基化研究中的应用

组成糖链的糖基种类繁多,连接方式多样,主要包括 *N*-连接糖链和 *O*-连接糖链。糖蛋白存在着复杂的微观不均一性(microheterogencity),即在蛋白质的氨基酸序列完全相同的情况下,由于糖链的组成或连接方式不同而形成多种糖蛋白亚型。糖蛋白的这种复杂性在其生理功能上具有重要意义,但却给糖蛋白的分析带来很多障碍。糖蛋白的分析主要包括几个层次。① 糖基化位点的确定。此处的困难在于有的糖蛋白的糖基化位点多,而且包括 *N*- 和 *O*- 糖基化位点,不容易采用单一方法确定。传统方法是通过 Edman 降解和糖苷酶酶解或化学降解相结合来确定位点,但对于多位点糖基化和不容易被糖苷酶去除的 *O*-糖链,此法难以奏效。② 糖链的组成,即糖链由哪些单糖组成。*N*-糖链有固定的五糖核心 $GlcNAc_2Man_3$,核心外再连接有不同长度的糖链,而 *O*-糖链则比较多样化。传统方法采用酶解或化学降解方法获得糖链后,水解成单糖,用阴离子交换色谱分离,然后与标准图谱对照。但这种方法仅能确定糖蛋白总的糖基组成,而无法获得某一位点上的糖链组成。③ 糖链不均一性的鉴定及各成分间的比例。困难在于难以分离到每

种成分并进行分析。④ 糖链的连接顺序。⑤ 糖链的分支和连接方式。⑥ 糖基异头构型的分析。

由于糖蛋白的分析困难,更需一种有效的分析手段。20 世纪 80 年代, Carr 和 Roberts 采用快原子轰击质谱分析了经 *N*-糖苷酶 F 处理前后的糖肽,确定了糖基化位点,观察到糖链的不同成分,部分序列及糖基化程度。采用这一方法分析了组织型纤溶酶原激活物的 *N*-糖链结构,还通过甲基化衍生对糖链连接方式进行了分析[71]。Sasaki 等则应用 HPLC,FAB-MS 和 Edman 降解结合的方法,鉴定了重组红细胞生成素的 3 个 *N*-糖链的组成,天线型和不均一性,以及一个 *O*-糖链的不均一组成[72]。20 世纪 90 年代初,电喷雾质谱(ESI-MS)和激光解吸-飞行时间质谱(MALDI-TOF)的应用,使糖蛋白的分析出现了突飞猛进的发展[73]。

1. 糖蛋白糖基化位点的分析

质谱法分析糖基化位点目前常采用特征离子监测 (selected ion monitor ,SIM)[74,75],以及蛋白质酶酶解和糖苷酶酶解相结合的方法。在 LC-ESI 质谱上,采用 CID 和特征离子监测(SIM)来寻找糖基特征离子,进而确定糖肽。如 *N*-乙酰葡糖胺 NacGlc 的特征离子在 m/z 204 ,乙酰唾液酸(NeuNAc) 的特征离子在 m/z 292。SIM 则可寻找到含这些糖特征离子的肽段。

但 SIM 方法有时也会有假阳性信号。因此,另一个方法是比较糖苷酶处理前后的蛋白质酶酶解肽谱[76]。在总离子流(TIC)图上,糖苷酶酶解前的糖肽(glycosylated peptide)峰在酶解后消失,而出现新的去糖肽(deglycosylated peptide)。根据去糖肽的相对分子质量和 MS/MS 可以确定肽段在蛋白质中的位置,以及糖基化位点。

2. 糖链的组成、连接顺序及连接方式分析

ESI 质谱和 MALDI-TOF 质谱在分析糖链方面各有优缺点。ESI 质谱较容易获得糖链结构信息,但糖链的离子化程度较差;MALDI-TOF 质谱适于检测糖链复杂的不均一性,但不适于分析细微的结构。

在 ESI 质谱中,糖链的组成和不均一性可以直接由糖肽的 MS 和 MS/MS 图谱解析,根据单糖的质量数可推断糖链的组成。但对于同质量异构的糖基(如 Man 和 Glu),仅从 MS 图谱还无法分辨。因此,通常采用专一的外切糖苷酶(如甘露糖苷酶、唾液酸酶等)处理糖肽后,再进行质谱分析,从而鉴定出糖链的组成和顺序[76]。

如果要进一步细致地分析糖链结构,一般是采用酶解或化学裂解的方法,将糖链从蛋白质上切割下来后进行 MS, MS/MS 乃至多级 MS 分析。但是从糖蛋白上释放出的寡糖,一般较难离子化。所以,在进行质谱分析时常常需要加以衍生,增强其疏水性和离子化能力[77]。常用的衍生化方法有甲基化和还原胺化等,具体的衍生试剂在本书上篇有详细介绍。

分析糖链结构,必然要进行碰撞诱导解离(CID),糖链在发生碰撞诱导解离时,有三种断裂方式。①单糖间相连键的断裂。由此获得糖链的线性序列 (linear sequence)。②单糖间相连键两次断裂。通过这种方式获得内部离子(internal ions),从而得到糖链的分支信息。③还原端交叉环断裂。它可以提供单糖间连接方式的信息[78]。但是,由于化学

键的强度不一，断裂方式①最容易发生，相应的碎片离子强度高，掩盖了由②、③两种断裂方式产生的碎片离子。因此，质谱法分析寡糖不太容易分辨糖链的分支和连接方式。

具有多级质谱功能的离子阱质谱和 FTICR-MS 在分析糖链结构方面有独特优势，Welskopf 等应用离子阱质谱的多级 MS(multistage MS, MSn)，分析了复杂 N 连接糖链，在多级 MS 中，获得了在二级 MS 中被掩盖或没有断裂的碎片离子，分析了糖链的分支及连接方式[79]。Cancilla 采用碱性降解和 MALDI-FT-MS 结合的方法，测定未经衍生的寡糖的序列，通过 MALDI-CID 或持续非共振照射 CID(SORI CID)获得了寡糖还原端交叉环(crossing-ring)断裂而成的碎片离子，成功测定了寡糖的连接方式[80]。

Q-TOF-MS 由于具有高分辨、高灵敏度和精确度，也正逐渐应用于糖链的研究，在灵敏度和结构解析方面都超过了三级四极杆质谱[77]。母离子扫描方法最初是基于三级四极杆的，随着 Q-TOF-MS 的出现，也开始在 Q-TOF-MS 上应用。与三级四极杆相比，Q-TOF-MS母离子扫描的分辨率更好，信噪比更高，但是灵敏度有所下降。母离子扫描(precursor ion scan)方法也开始在蛋白质糖基化研究中应用，通过扫描特征性的子离子(m/z)204($HexNAc^+$)、366($Hex\text{-}Hex\text{-}Nac^+$)等，找到其相应的母离子，并将这个母离子片段化，从而可以进行糖基化位点分析和糖链结构解析[81]。Q-TOF-MS 母离子扫描方法应用到蛋白质糖基化研究中，为质谱法分析糖蛋白提供了一个更好的手段。

MALDI-TOF-MS 由于可分析复杂混合物，所以能直接分析相对分子质量较大的糖蛋白，并分辨不均一性。MALDI-TOF-MS 也能方便地检测衍生或未经衍生的寡糖[82,83]。特别对于非衍生化的寡糖，MALDI-TOF-MS 比 ESI-MS 效果好。在具有 PSD 或 CID 功能的情况下，更能得到相当的结构信息[84~86]。

（六）糖基化蛋白质组学的研究策略

1. 凝胶电泳技术及质谱联用

传统的糖基化研究需要高纯度的蛋白质样品，而凝胶分离技术在糖基化研究中一个最明显的优势便是能分离混合物，并且对蛋白质的纯度没特殊要求。可针对不同的糖基连接方式，采取不同的途径进行分析。

(1) *N*-糖基化蛋白分析

混合物或经过凝集素预分级的糖蛋白混合物通过 SDS 聚丙烯酰胺凝胶电泳或双向凝胶电泳来分离蛋白，用免疫染色或凝集素染色从胶中找到感兴趣的条带或点。用糖苷酶(如肽 *N*-糖苷酶 F，即 PNGase F)胶内酶解释放出糖链，释放出的糖链可用荧光标记(如 2-氨基苯甲酰胺，即 2AB)、外切糖苷酶酶解以及高效液相色谱分析或用基质辅助激光解析离子化-飞行时间(MALDI-TOF)质谱来测定糖链相对分子质量，从而确定糖链结构[87,88]。用胰蛋白酶对胶内的蛋白质进行胶内酶解，抽提出肽段，用基质辅助激光解析离子化-飞行时间(MALDI-TOF)质谱或液相色谱-串联质谱(LC-MS/MS)等进行鉴定。此过程见图 18.17。

(2) *O*-糖基化蛋白分析

将蛋白质或其混合物经过 SDS-聚丙烯酰胺凝胶电泳(如胎牛球蛋白，Direct Blue 染色)或 SDS-琼脂糖/聚丙烯酰胺凝胶电泳(如黏蛋白，Alican Blue 染色)及双向凝胶电泳

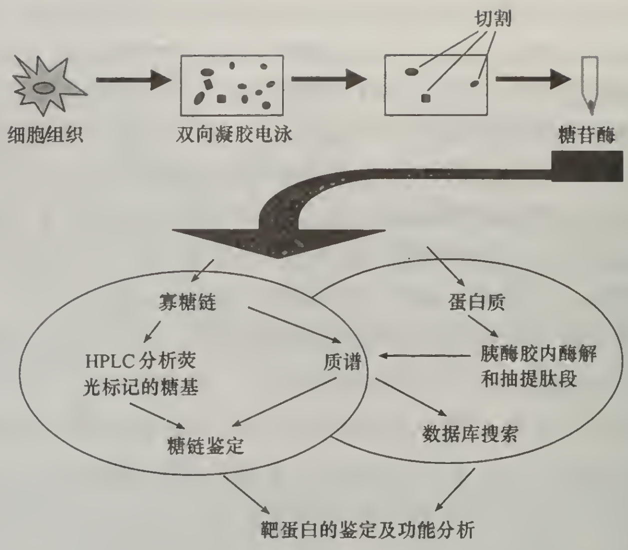


图 18.17　*N*-糖基化蛋白分析流程图

(如人血浆蛋白,Direct Blue 染色)后转移到 PVDF 膜,染色脱色后切取目标蛋白条带进行β-消除反应(甲醇浸湿后于 20μl 的 50mmol/L 氢氧化钠和 0.5mol/L 硼氢化钠溶液中,50℃温育 16h,1μl 冰乙酸中止反应),所得样品脱盐干燥后重溶于水中进行液相色谱/电喷雾离子化质谱分析[89]。

2. 非凝胶技术及质谱联用

蛋白质混合物经过凝集素琼脂糖柱 1 亲和层析,得到与该凝集素特异性结合的糖蛋白混合物;将糖蛋白混合物用胰蛋白酶或无色细菌(Achromobacter)蛋白酶Ⅰ(一种赖氨酸内肽酶)酶解,得到的肽段混合物,再经过与凝集素琼脂糖柱 1 相同的凝集素琼脂糖柱 1′(凝集素相同,柱体积可能不同)亲和层析。得到相应的糖肽混合物,再经过高效液相色谱进行分步收集。糖肽的肽段序列可通过测序仪或质谱等进行测序,一些糖链结构可通过质谱或前沿亲和层析法(FAC)等方法进行解析。未被凝集素琼脂糖柱 1 吸附的蛋白质混合物可经过凝集素琼脂糖柱 2 进行亲和层析,依此类推。

3. *O*-β-GlcNAc 糖基化的研究

O-GlcNAc 糖基化和去 *O*-GlcNAc 糖基化的酶(*O*-GlcNAc 转移酶和 *O*-GlcNAc 酶)已经被克隆和鉴定[90~92]。研究表明,*O*-GlcNAc 糖基化与磷酸化之间是一种"阴阳"关系[93]。我们一直认为磷酸化是调控蛋白质功能的基本机制,而很少有人意识到 *O*-GlcNAc 糖基化修饰的特殊作用。与磷酸化蛋白类似,*O*-GlcNAc 糖基化蛋白也能被筛选,并进行鉴定[94]。现在磷酸化研究已成为功能蛋白质组的一个研究热点,我们也应当关注如何更深入地研究蛋白质 *O*-GlcNAc 糖基化。应用蛋白质组的技术大规模研究蛋白质 *O*-GlcNAc 糖基化也势在必行。

在包含 *O*-GlcNAc 糖基化蛋白的蛋白质混合物中,用半乳糖基转移酶将氚(^{3}H)标记

的半乳糖(UDP-[6-^{3}H]-半乳糖)转移到 *O*-GlcNAc 上[95]。将标记好的蛋白质混合物用胰蛋白酶酶解,得到肽段混合物。再将肽段混合物经过蓖麻蛋白琼脂糖柱(RCA Ⅰ)进行亲和层析,得到标记了的肽段。这些肽段用微孔或毛细管液相色谱分离,用串联质谱进行检测。这里的质谱仪器是三级四极杆类型的,这样可以用较低的能量碰撞得到糖链离子(质荷比为 366),用质荷比 366 的母离子扫描可找到 *O*-GlcNAc 修饰的肽段,再将该肽段进行高能碰撞,得到的数据用 SEQUEST 软件进行数据库搜索,鉴定出这个被 *O*-GlcNAc 修饰的肽段。见图 18.18。

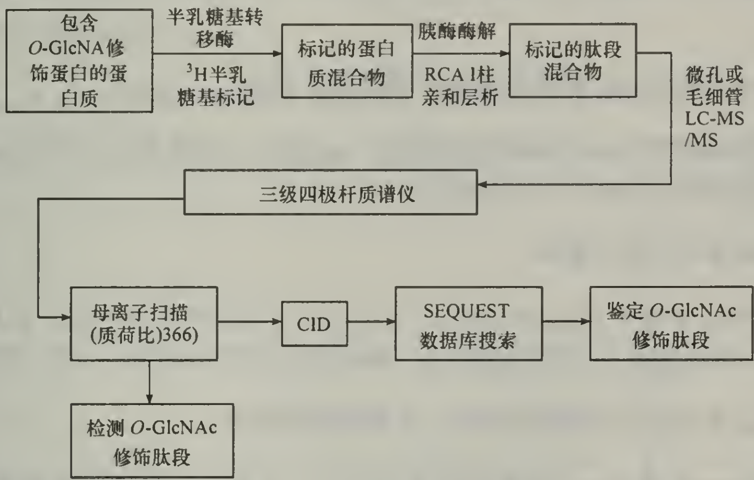


图 18.18　*O*-β-GlcNAc 糖基化分析的流程图

用抗 *O*-GlcNAc 的抗体柱纯化 *O*-GlcNAc 修饰的糖蛋白,在对 *O*-GlcNAc 进行 β-消除后,再通过 Michael 加成化学反应,用 DTT 或生物素戊胺替换糖苷键,因替换后键能稳固,适合用 CID MS/MS 的方法进行位点鉴定[96]。

二、材　　料

乙腈(ACN), HPLC 级, 为 Fluck 公司产品;三氟乙酸(TFA)为 Merck 公司产品;二硫苏糖醇(DTT)为 Calbiochem 公司产品;Glu-C 蛋白酶, *N*-糖苷酶 F, α-甘露糖苷酶为 Boehringer Mannheim 公司产品;碘代乙酰胺(ICH_2CONH_2),牛胰核糖核酸酶 B (RNase B), Lys-C 蛋白酶为 Sigma 公司产品。所用水为 Milli-Q 水。

三、方　　法

1. RNase B 完整相对分子质量的测定

RNase B 溶于水中至 1g/L,HPLC 上样 10μl。HPLC 梯度为 0～2min, 0%～0% Solvent B; 2～15min, 0%～100% Solvent B。

2．RNase B 的 Lys-C 酶解

RNase B 溶于 80μl 30mmol/L NH_4HCO_3（pH7.8）缓冲液中至最终浓度 2g/L，加入 50mmol/L DTT 10μl，50℃ 反应 30min，冷却至室温，加入 100mmol/L 碘代乙酰胺（ICH_2CONH_2）10μl，37℃反应 15min。反应后的甲基酰胺（S-carboxyamidomethyl）RNase B（CAM-RNase B）经 HPLC 纯化，溶于 50μLNH_4HCO_3（pH 7.8）中至最终浓度 1g/ L，加 0.5g/L 酶 3μl，37℃保温 4h，再补加酶 3μl，继续保温 18h。酶解完成后取出冷却至 −20℃ 保存。

3．RNase B 的 Lys-C 酶解产物的 *N*-糖苷酶 F 处理

CAM-RNase B 的 Lys-C 酶解后溶液中加入 1 单位 *N*-糖苷酶 F 水溶液 4μl，37℃ 反应 2h，反应完成后取出冷却至 −20℃ 保存。

4．RNase B 的 Glu-C 酶解

CAM-RNase B 溶于 50μl NH_4HCO_3（pH 7.8）中至最终浓度 1g/L，加 1g/L 酶 3μl，37℃保温 4h，再补加酶 3μl，继续保温 18h。酶解完成后取出冷却至 −20℃ 保存。

5．RNase B 的 Glu-C 酶解产物的 α-甘露糖苷酶处理

CAM-RNase B 的 Glu-C 酶解后溶液中加入 0.5 单位 α-甘露糖苷酶溶液（含 3.0 mol/L $(NH_4)_2SO_4$，0.1mmol/L 醋酸锌，pH 7.5）2μl，37℃ 反应 2h，反应完成后取出冷却至 −20℃ 保存。

6．LC-MS 测定条件

测定在 Finnigan MAT LCQ 离子阱质谱仪上进行，毛细管温度：180℃；喷雾电压：4.25 kV；鞘气流速：40；联用的 HPLC 为 HP-1100 型，反相 C8 柱（RP-300，30mm × 2.1mm I.D.，ABI 公司产品）或 C18 柱（150mm × 2.1mm I.D.，HP 公司产品）。流速：0.2ml/min；洗脱液：Solvent A，0.02% TFA-H_2O；Solvent B，2% H_2O-98% ACN-0.02% TFA。MS/MS 能量为 42%。HPLC 梯度为 0～2min 0%～0% B，2～90min 0%～50% B。

四、结果和讨论

1.RNase B 完整分子的分子质量测定

RNase B 由 124 个氨基酸组成，其序列为：

KETAAAKFERQHMDSSTSAASSSNYCNQMMKSRN* LTKDRCKPVNTFVHESL ADVQAVCSQKNVACKNGQTNCYQSYSTMSITDCRETGSSKYPNCAYKTTQANKHII VACEGNPYVPVHFDASV。

从氨基酸序列得出的理论平均分子质量为 13 682.4Da，但 34 位的 Asn 为糖基化位

点（用 N* 表示）。图 18.19（A）为 RNase B 的质荷比图，图 18.19（B）为 RNase B 的质量。两图都呈现明显的不均一型，而且分子质量也与由氨基酸序列得出的理论分子质量不同。实际得到 6 种不均一的蛋白质的质量，它们的氨基酸一级结构相同，造成不均一的原因是所连接的糖链长短不一。除了 5 种糖蛋白以外，也有不含糖链的蛋白质，测得分子质量为 13 683.7Da，这是由于含有核糖核酸酶 A，它的氨基酸序列与核糖核酸酶 B 相同，但不含糖链。5 种糖蛋白的分子质量分别为 14 901.8Da，15 064.4Da，15 227.3Da，15 386.3Da，15 549.1Da，两相邻的分子质量相差约 162Da，正好是一个六碳糖的分子质量。由此可验证糖链的不均一型。

2．RNase B Lys-C 酶解产物经 *N*-糖苷酶 F 处理前后的 LC-MS 分析

图 18.20（A）为 RNase B Lys-C 酶解产物经质谱分析的总离子流图。质谱测得图中

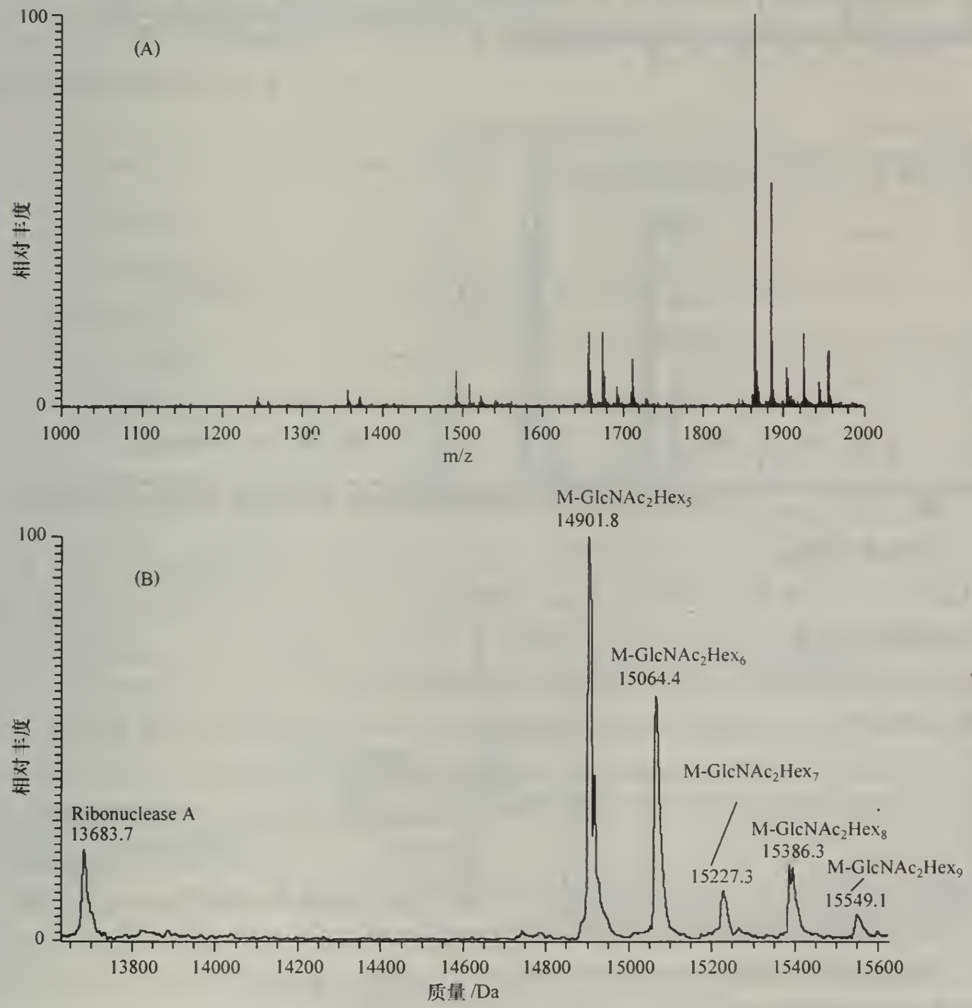


图 18.19　RNase B 完整分子的分子质量测定

（A）为质谱图，（B）为计算出的蛋白质分子质量

每个峰的分子质量,并确定了每个峰在蛋白质中的归属，但保留时间为 17～18min 一组峰的分子质量无法与氨基酸序列得到的理论分子质量匹配，推断是糖肽。图 18.20 (C) 为经 *N*-糖苷酶 F 处理的 RNase B Lys-C 酶解产物的总离子流图。所有肽的保留时间和分子质量都与处理前一致,而图 18.20 (A)中保留时间为 17～18min 的一组峰转化为在

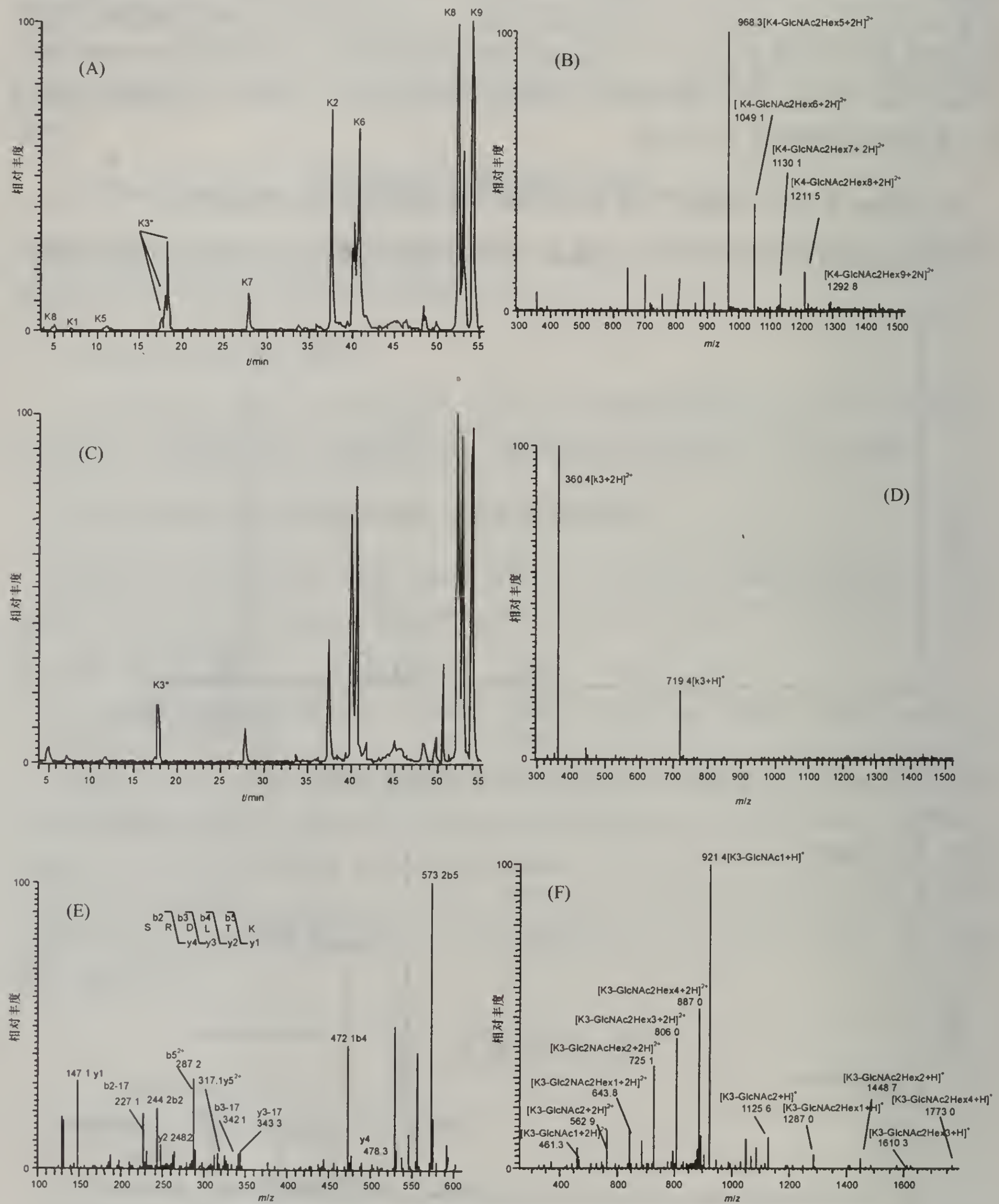


图 18.20　RNase B 的 Lys-C 酶解肽质量肽谱

(A)和(C)分别为糖苷酶-F 处理前后的酶解肽质谱总离子流图,K3*代表糖肽,K3$^{\Delta}$为糖苷酶-F 处理后去糖基化的肽;(D)和(E)分别为去糖基化肽的 MS 和 MS/MS (母离子为 m/z 360.4);(B)和(F)分别为糖肽的 MS 和 MS/MS(母离子为 m/z 968.3)

17min 处出现一个新的单一的峰，这个新峰的分子质量为 718.6Da，比肽段 32～38 的理论分子质量增加 1Da，为除去糖链的肽段 32～38。由于 *N*-糖苷酶 F 在去除 Asn34 上的糖链时，将 Asn（分子质量为 114.10Da）转化成 Asp（分子质量为 115.09Da），因此分子质量比原肽段 32～38 增加 1Da。图 18.20（D）为这段肽的质谱图，进一步 MS/MS 测得这段肽的序列，证实为肽段 32～38［图 18.20（E）］。因此，可以确认 RNase B Lys-C 酶解产物总离子流图中保留时间为 17～18min 的肽即为糖肽。图 18.20（B）为糖肽的 MS 图谱，同样呈现明显的不均一性。其中主要成分的分子质量为 1 934.3Da，比肽的分子质量增加 1 215.7Da，与糖链 $GlcNAc_2Hex_5$ 的分子质量相匹配，即在 *N*-连接糖链五糖核心外再连接两个六碳糖。MS/MS 结果得到了糖链的组成和连接顺序［图 18.20（F）］。除了主要成分外，还有连接有 $GlcNAc_2Hex_6$～$GlcNAc_2Man_3Hex_9$ 的另外 4 种成分。总共 5 种糖链类型造成了糖蛋白的微观不均一性，因此，在糖苷酶处理前糖肽呈现出一组峰，而处理后呈现为单一的峰。此外，由于核心外连接的都是六碳糖，糖型被推断为高甘露糖型。此外，我们还对不均一酶解糖肽进行了相对定量，根据质谱测定的离子强度，确定了含有不同长短糖链糖肽的相对含量（表 18.1）。

表 18.1　不同糖型的相对定量

	胰酶酶解后的糖肽组成百分比
K3-GlcNAc2Man9	47.7%
K3-GlcNAc2Man8	23.8%
K3-GlcNAc2Man7	8.9%
K3-GlcNAc2Man6	15.8%
K3-GlcNAc2Man5	3.8%

注：GlcNAc 为 *N*-乙酰葡糖胺；Man 为甘露糖。

3．RNase B Glu-C 酶解产物 α-甘露糖苷酶处理前后 LC-MS 分析

为了进一步证实糖链为高甘露糖型，我们对 CAM-RNase B Glu-C 酶解产物进行 α-甘露糖苷酶处理，分别经 0，5，22，31，50h 处理，取样由 LC-MS 分析。图 18.21（A）和（B）分别为 CAM-RNase B Glu-C 酶解产物未经 α-甘露糖苷酶处理的总离子流图和糖肽的 MS 图谱。α-甘露糖苷酶可以依次去除高甘露糖型糖链中五糖核心外的所有甘露糖和核心内的两个甘露糖，但不能去除核心内与 *N*-乙酰葡糖胺连接的 β-甘露糖。从不同时间取样的 MS 结果可看出，α-甘露糖苷酶依次去除了甘露糖。图 18.21（C）和（D）为处理 50h 后的总离子流图和糖肽的 MS 图谱。糖蛋白连接的糖型确定为高甘露糖型，糖链结构为 Glc-NAc2Man5-GlcNAc2Man9。

我们采用液相色谱电喷雾质谱，结合蛋白酶解和糖苷酶解的方法，为糖蛋白的分析提供了快速、灵敏的检测手段。通常，对糖蛋白上糖链的结构分析需要将糖链切割下来并回收，再进行色谱分析或质谱分析。我们直接比较糖苷酶处理前后的糖蛋白酶解肽谱，可以确定糖肽的位置、糖结合位点，并通过串联质谱分析糖肽上所连糖链的结构。对于高甘露糖型的糖链，我们采用酶解肽段经不同时间的 α-甘露糖苷酶处理，取样进行质谱分析的方法，可以区别高甘露糖型和其他 *N*-连接的糖链类型。我们采用在糖蛋白酶解和糖苷酶酶

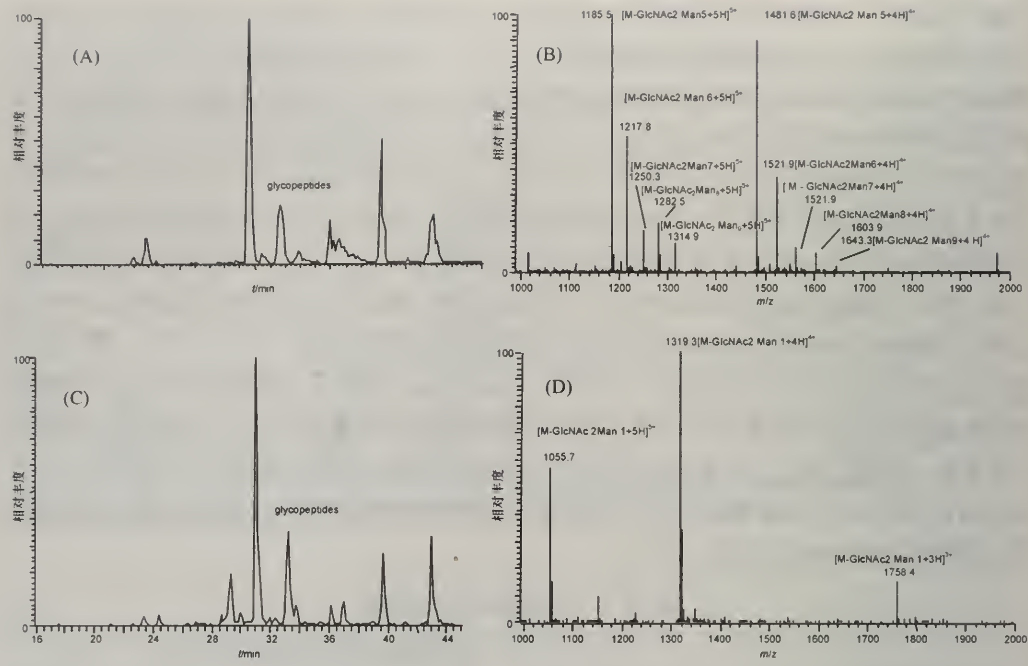


图 18.21　RNase B 的 Glu-C 酶解肽质量肽谱

(A)和(C)分别为 α-半乳糖苷酶处理前后的酶解肽质谱总离子流图；

(B)和(D)分别为 α-半乳糖苷酶处理前后的糖肽和去糖基化肽的质谱图

解连续进行，不必再进行分离纯化。只是要求蛋白质酶解和糖苷酶酶解的 pH 相匹配，如 *N*-糖苷酶 F 一般在 pH7.5 左右使用，与 Lys-C 匹配，而 α-甘露糖苷酶的最适 pH 在 4.0～5.0，与 Glu-C 可匹配使用。采用这种方法，不同酶解的图谱总离子流图重复性很好，便于比较糖苷酶酶解处理前后的图谱，寻找糖肽或糖链有变化的肽段。这一方法尤其适用于具有 *N*-糖链的糖蛋白分析，因为 *N*-糖链对糖苷酶较敏感，而某些 *O*-糖链不容易通过酶解方法完全去除。

五、展　　望

基于糖蛋白本身的复杂性，我们在蛋白质糖基化研究中必须运用多元化的策略和技术。Fukui 等人开始用寡糖微阵列的方法来研究糖和蛋白质的相互作用[97]。现在可在 LCQ DecaXP(Thermo Finnigan)的离子阱中做四级质谱(MS4)来解析糖肽的结构[98]。美国新罕布什大学化学系的结构生物学研究中心现在开始用基质辅助激光解析离子化-四极杆离子阱-飞行时间(MALDI-quadrupole ion trap-TOF, Shimadzu Biotech)进行线虫(*C. elegans*)糖基化蛋白质组学的研究。随着分离技术、检测技术的迅速发展，蛋白质糖基化研究的方法和策略也将呈现出多样性，糖基化蛋白质组研究也将成为翻译后修饰研究中的一个新热点。

（阮宏强　周　虎　曾　嵘）

参 考 文 献

1 Krebs EG. Historical perspectives on protein phosphorylation and a classification system for protein kinases. Philos Trans R Soc Lond B Biol Sci ,1983,302: 3～11

2 Yan JX, Packer NH, Gooley AA *et al*. Protein phosphorylation: technologies for the identification of phosphoamino acids. J Chromatogr A,1998,808: 23～41

3 Mann M, Ong SE, Gronborg M *et al*. Analysis of protein phosphorylation using mass spectrometry: deciphering the phosphoproteome. Trends Biotechnol,2002,20: 261～268

4 Zolnierowicz S, Bollen M. Protein phosphorylation and protein phosphatases. De Panne, Belgium, September 19～24, 1999. EMBO J,2000,19: 483～488

5 Venter JC, Adams MD, Myers EW *et al*. The sequence of the human genome. Science, 2001,291:1304～1351

6 Gronborg M, Kristiansen TZ, Stensballe A *et al*. A mass spectrometry-based proteomic approach for identification of serine/threonine-phosphorylated proteins by enrichment with phospho-specific antibodies: identification of a novel protein, frigg, as a protein kinase a substrate. Mol Cell Proteomics,2002,1: 517～527

7 McLachlin DT, Chait BT. Analysis of phosphorylated proteins and peptides by mass spectrometry. Curr Opin Chem Biol,2001,5: 591～602

8 Stensballe A, Jensen ON, Olsen JV *et al*. Electron capture dissociation of singly and multiply phosphorylated peptides. Rapid Commun. Mass Spectrom,2000,14: 1793～1800

9 Pandey A, Podtelejnikov AV, Blagoev B *et al*. Analysis of receptor signaling pathways by mass spectrometry: identification of vav-2 as a substrate of the epidermal and platelet-derived growth factor receptors. Proc Natl Acad Sci USA,2000, 97: 179～184

10 Zhou H, Watts JD, Aebersold R *et al*. A systematic approach to the analysis of protein phosphorylation. Nat Biotechnol,2001,19: 375～378

11 Oda Y, Nagasu T, Chait BT *et al*. Enrichment analysis of phosphorylated proteins as a tool for probing the phosphoproteome. Nat Biotechnol,2001,19: 379～382

12 Goshe MB, Conrads TP, Panisko EA *et al*. Phosphoprotein isotope-coded affinity tag approach for isolating and quantitating phosphopeptides in proteome-wide analyses. Anal Chem,2001,73: 2578～2586

13 Jaffe H, Veeranna, Pant HC *et al*. Characterization of serine and threonine phosphorylation sites in beta-elimination/ethanethiol addition-modified proteins by electrospray tandem mass spectrometry and database searching. Biochemistry, 1998,37: 16211～16224

14 Raska CS, Parker CE, Dominski Z *et al*. Direct MALDI-MS/MS of phosphopeptides affinity-bound to immobilized metal ion affinity chromatography beads. Anal Chem,2002,74: 3429～3433

15 Hart SR, Waterfield MD, Burlingame AL *et al*. Factors governing the solubilization of phosphopeptides retained on ferric NTA IMAC beads and their analysis by MALDI TOFMS. J Am Soc Mass Spectrom,2002,13: 1042～1051

16 Posewitz MC, Tempst P. Immobilized gallium(Ⅲ) affinity chromatography of phosphopeptides. Anal Chem,1999,71: 2883～2892

17 Cao P, Stults JT. Mapping the phosphorylation sites of proteins using on-line immobilized metal affinity chromatography/capillary electrophoresis/electrospray ionization multiple stage tandem mass spectrometry. Rapid Commun Mass Spectrom, 2000,14: 1600～1606

18 Stensballe A, Andersen S, Jensen ON *et al*. Characterization of phosphoproteins from electrophoretic gels by nanoscale Fe(Ⅲ) affinity chromatography with off-line mass spectrometry analysis. Proteomics,2001,1: 207～222

19 Xhou W, Merrick BA, Khaledi MG *et al*. Detection and sequencing of phosphopeptides affinity bound to immobilized metal ion beads by matrix-assisted laser desorption/ionization mass spectrometry. J Am Soc Mass Spectrom,2000,11: 273～282

20 Ficarro SB, McCleland ML, Stukenberg PT *et al*. Phosphoproteome analysis by mass spectrometry and its application to Saccharomyces cerevisiae. Nat Biotechnol,2002,20: 301～305

21 Yip TT, Hutchens TW. Mapping and sequence-specific identification of phosphopeptides in unfractionated protein digest mixtures by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. FEBS Lett, 1992, 308: 149～153

22 Liao PC, Leykam J, Andrews PC *et al*. An approach to locate phosphorylation sites in a phosphoprotein: mass mapping by combining specific enzymatic degradation with matrix-assisted laser desorption/ionization mass spectrometry. Anal Biochem, 1994, 219: 9～20

23 Amankwa LN, Harder K, Jirik F *et al*. High-sensitivity determination of tyrosine-phosphorylated peptides by on-line enzyme reactor and electrospray ionization mass spectrometry. Protein Sci, 1995, 4: 113～125

24 Jonscher KR, Yates JR Ⅲ. Matrix-assisted laser desorption ionization/quadrupole ion trap mass spectrometry of peptides. Application to the localization of phosphorylation sites on the P protein from Sendai virus. J Biol Chem, 1997, 272: 1735～1741

25 Qin J, Chait BT. Identification and characterization of posttranslational modifications of proteins by MALDI ion trap mass spectrometry. Anal Chem, 1997, 69: 4002～4009

26 Moyer SC, Cotter RJ, Woods AS *et al*. Fragmentation of phosphopeptides by atmospheric pressure MALDI and ESI/Ion trap mass spectrometry. J Am Soc Mass Spectrom, 2002, 13: 274～283

27 Reid GE, Simpson RJ, O'Hari RA *et al*. Leaving group and gas phase neighboring group effects in the side chain losses from protonated serine and its derivatives. J Am Soc Mass Spectrom, 2000, 11: 1047～1060

28 Zappacosta F, Huddleston MJ, Karcher RL *et al*. Improved sensitivity for phosphopeptide mapping using capillary column HPLC and microionspray mass spectrometry: comparative phosphorylation site mapping from gel-derived proteins. Anal Chem, 2002, 74: 3221～3231

29 Tholey A, Reed J, Lehmann WD *et al*. Electrospray tandem mass spectrometric studies of phosphopeptides and phosphopeptide analogues. J Mass Spectrom, 1999, 34: 117～123

30 Watty A, Neubauer G, Dreger M *et al*. The *in vitro* and *in vivo* phosphotyrosine map of activated MuSK. Proc Natl Acad Sci USA, 2000, 97: 4585～4590

31 Knotts TA, Orkiszewski RS, Cook RG *et al*. Identification of a phosphorylation site in the hinge region of the human progesterone receptor and additional amino-terminal phosphorylation sites. J Biol Chem, 2001, 276: 8475～8483

32 Annan RS, Huddleston MJ, Verma R *et al*. A multidimensional electrospray MS-based approach to phosphopeptide mapping. Anal Chem, 2001, 73: 393～404

33 Steen H, Mann M. A new derivatization strategy for the analysis of phosphopeptides by precursor ion scanning in positive ion mode. J Am Soc Mass Spectrom, 2002, 1:996～1003

34 Bateman RH, Carruthers R, Hoyes JB *et al*. A novel precursor ion discovery method on a hybrid quadrupole orthogonal acceleration time-of-flight (Q-TOF) mass spectrometer for studying protein phosphorylation. J Am Soc Mass Spectrom, 2002, 13: 792～803

35 Steen H, Kuster B, Fernandez M *et al*. Detection of tyrosine phosphorylated peptides by precursor ion scanning quadrupole TOF mass spectrometry in positive ion mode. Anal Chem, 2001, 73:1440～1448

36 Hoffmann R, Wachs WO, Berger RG *et al*. Chemical phosphorylation of the peptides GGXA (X = S, T, Y): an evaluation of different chemical approaches. Int J Pept Protein Res, 1995, 45: 26～34

37 Hunter AP, Games DE. Chromatographic and mass spectrometric methods for the identification of phosphorylation sites in phosphoproteins. Rapid Commun Mass Spectrom, 1994, 8: 559～570

38 Schlosser A, Pipkorn R, Bossemeyer D *et al*. Analysis of protein phosphorylation by a combination of elastase digestion and neutral loss tandem mass spectrometry. Anal Chem, 2001, 73: 170～176

39 Sadagopan N, Malone M, Watson JT *et al*. Effect of charge derivatization in the determination of phosphorylation sites in peptides by electrospray ionization collision-activated dissociation tandem mass spectrometry. J Mass Spectrom, 1999, 34: 1279～1282

40 Zarling AL Ficarro SB, White FM *et al*. Phosphorylated peptides are naturally processed and presented by major histocompatibility complex class Ⅰ molecules *in vivo*. J Exp Med, 2000, 192: 1755～1762

41 Li W, Boykins RA, Backlund PS *et al*. Identification of phosphoserine and phosphothreonine as cysteic acid and beta-methylcysteic acid residues in peptides by tandem mass spectrometric sequencing. Anal Chem, 2002, 74: 5701～5710
42 Adam GC, Sorensen EJ, Cravatt BF *et al*. Chemical strategies for functional proteomics. Mol Cell Proteomics, 2002, 1: 781～790
43 McLafferty FW, Horn DM, Breuker K *et al*. Electron capture dissociation of gaseous multiply charged ions by Fourier-transform ion cyclotron resonance. J Am Soc Mass Spectrom, 2001, 12: 245～249
44 Shi SD, Hemling ME, Carr SA *et al*. Phosphopeptide/phosphoprotein mapping by electron capture dissociation mass spectrometry. Anal Chem, 2001, 73: 19～22
45 Tsay YG, Wang YH, Chiu CM *et al*. A strategy for identification and quantitation of phosphopeptides by liquid chromatography/tandem mass spectrometry. Anal Biochem, 2000, 287: 55～64
46 Vener AV, Harms A , Sussman MR *et al*. Mass spectrometric resolution of reversible protein phosphorylation in photosynthetic membranes of Arabidopsis thaliana. J Biol Chem, 2001, 276: 6959～6966
47 Oda Y, Huang K, Cross FR *et al*. Accurate quantitation of protein expression and site-specific phosphorylation. Proc Natl Acad Sci USA, 1999, 96: 6591～6596
48 Goshe MB, Conrads TP, Panisko EA *et al*. Phosphoprotein isotope-coded affinity tag approach for isolating and quantitating phosphopeptides in proteome-wide analyses. Anal Chem, 2001, 73: 2578～2586
49 Weckwerth W, Willmitzer L, Fiehn O *et al*. Comparative quantification and identification of phosphoproteins using stable isotope labeling and liquid chromatography/mass spectrometry. Rapid Commun Mass Spectrom, 2000, 14: 1677～1681
50 Ruse CI, Willard B, Jin JP *et al*. Quantitative dynamics of site-specific protein phosphorylation determined using liquid chromatography electrospray ionization mass spectrometry. Anal Chem, 2002, 74: 1658～1664
51 Rademacher TW, Parekh RB, Dwek RA. Glycobiology. Annual Review of Biochemistry, 1988, 57:785～838
52 糖链结构与功能调控前瞻．香山科学会议第 192 次学术讨论会会议公告，2002 年 10 月 9 日至 11 日。
53 Hirabayashi J, Kasai K. Separation technologies for glycomics. J Chromatogr B Analyt Technol Biomed Life Sci, 2002, 771(1-2):67～87
54 Varki A, Cummings R, Esko J *et al*. Essentials of glycobiology 糖生物学．清华大学出版社 & 美国冷泉港实验室出版社，2002. 1～3, 10～11
55 Apweiler R, Hermjakob H, Sharon N. On the frequency of protein glycosylation, as deduced from analysis of the SWISS-PROT database. Biochim Biophys Acta, 1999, 1473:4～8
56 Laine RA. A calculation of all possible oligosaccharide isomers both branched and linear yields 1.05×10^{12} structures for a reducing hexasaccharide: the Isomer Barrier to development of single-method saccharide sequencing or synthesis system. Glycobiology, 1994, 4:759～767
57 Hirabayashi J, Kasai K. Glycomics, coming of age! Trends Glycosci Glycotechnol, 2000, 12(1):1～5
58 蒋燕灵，王克夷译．正在来临的糖组学．生命的化学，2000, 20(6):286～287
59 Wells L, Vosseller K, Hart GW. Glycosylation of nucleocytoplasmic proteins: signal transduction and O-GlcNAc. Science, 2001, 291(5512):2376～2378
60 Jung E, Veuthey A, Gasteiger E *et al*. Annotation of glycoproteins in the SWISS-PROT database. Proteomics ,2001, 1: 262～268
61 Gerrit J, Gerwing and Johannes FG Vliegenthart. Analysis of glycoprotein-derived glycopeptides. Proteomics in Function Genomics, Birkhauser, Jolles, Pierre, Jornvall, Hans, 2000, 1:159～185
62 Kasai K, Oda Y, Nishikawa S *et al*. Theory for its application to studies on specific interaction of biomolecules. J Chromatogr, 1986, 376:33～47
63 Hirabayashi J, Arata Y, Kasai Y. Reinforcement of frontal affinity chromatography for effective analysis of lectin-oligosaccharide interactions. J Chromatogr A, 2000, 890:261～271
64 Wayne F, Patton. Detection technologies in proteome analysis. J Chromatogr B, 2002, 771: 3～31
65 Eckhardt AE, Hayes CE, Goldstein IJ. A sensitive fluorescent method for the detection of glycoproteins in polyacry-

lamide gels. Anal Biochem, 1976, 73(1): 192～197
66 Che FY, Song JF, Yu LR *et al*. Fluorescent staining of glycoproteins on polyvinylidene difluoride membrane with 8-aminonaphthalene-1, 3, 6-trisulfonate. Biotechniques, 2001, 30(6): 1272～1278
67 Steinberg TH, Pretty *On Top* K, Berggren KN *et al*. Rapid and simple single nanogram detection of glycoproteins in polyacrylamide gels and on electroblots. Proteomics, 2001, 1(7): 841～855
68 刘厚宝,王炳生,陈惠黎等．胆管癌胆汁糖蛋白及其糖链结构的初步研究．上海医科大学学报,2000,27(5): 364～367
69 金晓凌,王炳生,陈惠黎等．凝集素点印迹法分析胆汁糖蛋白糖链结构诊断胆管癌．实用癌症杂志,2001,16(2): 178～180
70 Anumula KR, Du P. Characterization of carbohydrates using highly fluorescent 2-aminobenzoic acid tag following gel electrophoresis of glycoproteins. Anal Biochem, 1999, 275: 236～242
71 Carr SA, Roberts GD, Jurewicz A *et al*. Structural fingerprinting of Asn-linked carbohydrates from specific attachment sites in glycoproteins by mass spectrometry: application to tissue plasminogen activator. Biochimie, 1988, 70(11): 1445～1454
72 Sasaki H, Ochi N, Dell A *et al*. Site-specific glycosylation of human recombinant erythropoietin: analysis of glycopeptides or peptides at each glycosylation site by fast atom bombardment mass spectrometry. Biochemistry, 1988, 27(23): 8618～8626
73 Harvey DJ. The role of mass spectrometry in glycobiology. Glycoconj J, 1992, 9(1): 1～12
74 Schindler PA, Settineri CA, Collet X *et al*. Site-specific detection and structural characterization of the glycosylation of human plasma proteins lecithin: cholesterol acyltransferase and apolipoprotein D using HPLC/electrospray mass spectrometry and sequential glycosidase digestion. Protein Sci, 1995, 4: 791～803
75 Carr SA, Huddleston MJ, Bean MF. Selective identification and differentiation of N- and O-linked oligosaccharides in glycoproteins by liquid chromatography-mass spectrometry. Protein Sci, 1993, 2: 183～196
76 Zeng R, Xu Q, Shao XX *et al*. Characterization and analysis of a novel glycoprotein from snake venom using liquid chromatography-electrospray mass spectrometry and Edman degradation. Eur J Biochem, 1999, 266: 352～358
77 Dell A, Morris HR. Glycoprotein structure determination by mass spectrometry. Science, 2001, 291: 2351～2356
78 Sasaki H, Ochi N, Dell A *et al*. Site-specific glycosylation of human recombinant erythropoietin: analysis of glycopeptides or peptides at each glycosylation site by fast atom bombardment mass spectrometry. Biochemistry, 1988, 27: 8618～8626
79 Weiskopf AS, Vouros P, Harvey DJ. Electrospray ionization-ion trap mass spectrometry for structural analysis of complex N-linked glycoprotein oligosaccharides. Anal Chem, 1998, 70: 4441～4447
80 Cancilla MT, Penn SG, Lebrilla CB. Alkaline degradation of oligosaccharides coupled with matrix-assisted laser desorption/ionization Fourier transform mass spectrometry: a method for sequencing oligosaccharides. Anal Chem, 1998, 70: 663～672
81 Mark AR, Andrew CG, Michael JD *et al*. Precursor Ion Scanning for Detection and Structural Characterization of Heterogeneous Glycopeptide Mixtures. J Am Soc Mass Spectrom, 2002, 13: 1065～1077
82 Mechref Y, Baker AG, Novotny MV. Matrix-assisted laser desorption/ionization mass spectrometry of neutral and acidic oligosaccharides with collision-induced dissociation. Carbohydr Res, 1998, 313: 145～155
83 Okamoto M, Takahashi K, Doi T *et al*. High-sensitivity detection and postsource decay of 2-aminopyridine-derivatized oligosaccharides with matrix-assisted laser desorption/ionization mass spectrometry. Anal Chem, 1997, 69: 2919～2926
84 Harvey DJ. Matrix-assisted laser desorption/ionization mass spectrometry of carbohydrates. Mass Spectrom Rev, 1999, 18: 349～450
85 Harvey DJ, Bateman RH, Green MR. High-energy collision-induced fragmentation of complex oligosaccharides ionized by matrix-assisted laser desorption/ionization mass spectrometry. J Mass Spectrom, 1997, 32: 167～187
86 Yamagaki T, Nakanishi H. Ion intensity analysis of post-source decay fragmentation in curved-field reflectron matrix-assisted laser desorption/ionization time-of-flight mass spectrometry of carbohydrates: for structural characterization of

glycosylation in proteome analysis. Proteomics, 2001, 1:329～339

87 Rudd PM, Geoffrey RG, Bernhard K *et al*. Oligosaccharide sequencing technology. Nature, 1997, 388:205～208

88 Rudd PM, Colominas C, Royle L *et al*. A high-performance liquid chromatography based strategy for rapid, sensitive sequencing of N-linked oligosaccharide modifications to proteins in sodium dodecyl sulphate polyacrylamide electrophoresis gel bands. Proteomics, 2001, 1(2):285～294

89 Schulz BL, Packer NH, Karlsson NG. Small-Scale analysis of O-linked oligosaccharides from glycoproteins and mucins separated by gel electrophoresis. Anal Chem, 2002, 74:6088～6097

90 Kreppel LK, Blomberg MA, Hart GW. Dynamic glycosylation of nuclear and cytosolic proteins: cloning and characterization of a unique o-glcnac transferase with multiple tetratricopeptide repeats. J Biol Chem, 1997, 272:9308～9315

91 Lubas WA, Frank DW, Krause M *et al*. O-linked GlcNAc transferase is a conserved nucleocytoplasmic protein containing tetratricopeptide repeats. J Biol Chem, 1997, 272:9316～9324

92 Gao Y, Wells L, Comer FI *et al*. Dynamic O-glycosylation of nuclear and cytoslic proteins: cloning and characterization of a neutral, cytosolic -N-acetylglucosaminidase from human brain. J Biol Chem, 2001, 276:9838～9845

93 Hart GW, Gries KD, Dong LY *et al*. O-linked N-acetylglucosamine: the yin-yang' of Ser/Thr phosphorylation? Nuclear and cytoplasmic glycosylation. Adv Exp Med Biol, 1995, 376:115～123

94 Vosseller K, Wells L, Hart GW. Nucleocytoplasmic O-glycosylation: O-GlcNAc and functional proteomics. Biochimie, 2001, 83(7):575～581

95 Haynes PA, Aebersold R. Simultaneous detection and identification of O-GlcNAc-modified glycoproteins using liquid chromatography-tandem mass spectrometry. Anal Chem, 2000, 72:5402～5410

96 Wells L, Vosseller K, Cole RN *et al*. Mapping sites of O-GlcNAc modification using affinity tags for serine and threonine post-translational modifications. Mol Cell Proteomics, 2002, 1(10):791～804

97 Fukui S, Feizi T, Galustian C *et al*. Oligosaccharide microarrays for high-throughput detection and specificity assignments of carbohydrate-protein interactions. Nat Biotechnol, 2002, 20:1011～1017

98 Shiaw-Lin W, Bondarenko P, Tom S *et al*. Structural analysis of glycosylated peptides in complex mixtures with ion trap MSn. Application Report No. 300

第十九章　亚细胞蛋白质组学

第一节　概　　论

蛋白质组学研究通过蛋白质分离技术、鉴定技术以及生物信息学的结合来鉴别复杂生物结构的所有蛋白质组成。蛋白质组已被理解为一个基因组或一个生物体所表达的所有蛋白质，或者在多细胞生物体中一个组织或一类分化细胞所表达的所有蛋白质组成[1]。双向凝胶电泳技术(two-dimensional gel electrophoresis，2-DE)分离蛋白质混合物为在分子水平上研究生物复合体提供了强有力的工具[2~4]。多肽质谱指纹技术(peptide mass fingerprinting，PMF)[5~9]与自动化蛋白质数据库检索功能[10]的联合应用使高通量鉴定从 2-DE 胶上切下的蛋白质点得以实现。在过去不足 10 年时间内，已经建立了许多来自不同组织、体液和细胞株等样品的 2-DE 参考蛋白质组图谱(http://www.expasy.ch/ch2d/2dindex.html)[11]。

但真核生物体是一个复杂系统。以人为例，基因组分析研究表明，人具有 3 万～5 万个基因，经 mRNA 不同的剪切方式以及翻译后修饰将会产生更多的蛋白质种类。而且，受蛋白质等电点极限、疏水性、相对分子质量范围及蛋白质丰度限制，很难由 2-DE 技术一步到位地获得生物体的全蛋白质组[12,13]。因此，在用 2-DE 技术展示蛋白质点分布之前，引进新的分离措施是必要的。

随着多物种基因组计划的实施，在后基因组时代对基因序列表达产物的功能分析将日益重要。蛋白质的亚细胞定位对蛋白质生物功能的发挥起着极关键的作用，蛋白质亚细胞定位信息强烈暗示着其生物学功能。蛋白质亚细胞分级技术根据理化性质的不同分离纯化出各种亚细胞组分，极大地降低了针对全细胞蛋白质组学研究的复杂性。而且，对亚细胞组分所有蛋白质的鉴定，一方面可以明确各种亚细胞组分的蛋白质组成，另一方面可以获得相应蛋白质的亚细胞定位信息。对一些特征性亚细胞组分(如各种细胞器等)所有蛋白质组成的了解可促进对其结构与功能的理解；对新蛋白质亚细胞定位信息的了解可促进对其生物学功能的认识。亚细胞蛋白质组学(subcellular proteomics)已日益成为蛋白质组学研究的一个焦点[14]。

一、亚细胞蛋白质组学内涵

所谓亚细胞蛋白质组学，是针对细胞内不同区域结构功能单位的蛋白质组学研究。细胞内成分根据空间结构、分布以及功能不同，分成不同细胞器或细胞区域，如细胞膜、胞浆、线粒体、溶酶体、过氧化物酶体、内质网、高尔基体和细胞核等。另外，细胞器内一些功能单位多是大分子结构体(macromolecular structure)或多蛋白质复合体(multiprotein complex)，如核基质(nuclear matrix)、剪接体(spliceosome)、纺锤体(spindle pole)、核孔结构(nuclear pore complex)以及核糖体(ribosome)等[15]。这些细胞器、多蛋白质组成的功能

单位的区域化(compartmentalization)分布和聚集可被看作是蛋白质除等电点、相对分子质量之外的第三向分离参数而用于对全细胞蛋白质进行分离,对其进行蛋白质组学研究可以更好地理解全细胞的蛋白质组。最近有报道采用免疫共沉淀或亲和层析的方法分离出细胞或细胞器内一类信号复合体(signaling complexe)[16~18]或一类特性的蛋白质群体[19]进行蛋白质组学研究,更丰富了亚细胞蛋白质组学的内涵。

二、亚细胞蛋白质组学的意义

亚细胞分离技术本身在细胞与分子生物学研究中的重要作用已不言而喻。以亚细胞分级与蛋白质组学技术相结合为基础的亚细胞蛋白质组学在细胞与分子生物学研究中具有更为重要的意义。

1. 富集低丰度蛋白质,完善全细胞蛋白质组

由于全细胞表达的蛋白质种类繁多,直接做全细胞的蛋白质组学分析,受上样量及蛋白质组学技术本身的限制,难以分离、检测到全部蛋白质,尤其是一些低丰度蛋白质。但各种蛋白质在细胞内有一定的空间分布,各种亚细胞组分也有各自特征性的蛋白质种类组成。分离亚细胞组分可以减少蛋白质种类、富集低丰度蛋白质,有望分离鉴定出更多的新蛋白质、新基因。如 Fialka 等人[20]对 MDCK Ⅱ 细胞的内吞小泡(endocytic vesicles)的蛋白质组学研究发现了 9 个新蛋白质并鉴定出其中一个。不同亚细胞蛋白质组的汇合,可以完善全细胞蛋白质组。

2. 提供蛋白质的亚细胞定位信息

成熟蛋白质必须分拣到特定的细胞部位才能发挥其生物学功能,蛋白质的亚细胞定位信息会强烈暗示其生物学功能。目前各种蛋白质数据库中蛋白质的亚细胞定位信息很不完善。Jung 等[14]利用 PSORT 软件通过分拣信号(sorting signal)的检测将 SWISS-PROT、TrEMBL、TrEMBL-NEW 蛋白质数据库中的共 24 846 个人类蛋白质进行了亚细胞定位预测,发现其中约有 3 362 个线粒体蛋白质、12 170 个细胞核蛋白质以及 3 791 个内质网或高尔基体蛋白质等。这些预测需要实验证实,而且仍有许多蛋白质的亚细胞定位未能预测。另外,人类基因组测序分析表明人具有 3 万~5 万个基因,经 mRNA 不同的剪切方式以及翻译后修饰将会产生更多的蛋白质种类。这预示还有更多的新蛋白质待发现与定位。亚细胞蛋白质组学研究可直接提供大量蛋白质的亚细胞定位信息,促进蛋白质数据库的日趋完善。

3. 加深对亚细胞组分的结构功能理解

对特定亚细胞组分(如细胞器或多蛋白质功能复合体)的所有蛋白质组成的了解非常有利于针对该结构功能单位的特别研究。如 Rout 等人[21]从分离出的酵母核孔复合体(nuclear pore complex)组分中鉴定出 174 种蛋白质,并证明了其中 40 种蛋白质与核孔复合体有关,这对核孔复合体的结构模型与功能研究有积极意义。

4. 加深对细胞生理分子机制的理解

亚细胞分离技术的进展已能分离出更多的亚细胞多蛋白质结构功能单位。对核内剪接体(spliceosome complex)、纺锤体(spindle pole)以及信号复合体(signaling complex)等的蛋白质组学研究,促进了对细胞生理的分子生物学机制的探索。Wigge 等人[22]用免疫电镜方法观察到从酵母纺锤体组分中分离到的 11 种新蛋白质中的 8 种共定位于纺锤体上。Heller 等人[17]分离与 CD28 相关的信号复合体,鉴定出 4 种新蛋白质与 CD28 作用,提示了 CD28 分子在 T 细胞信号传导中的一种新的可能机制。

5. 亚细胞蛋白质组的差异表达谱

病理情况下,不同亚细胞组分受影响程度不一。适宜的亚细胞蛋白质组学研究将会更灵敏地反映病理因素造成的一些蛋白质的量变、质变及蛋白质迁移情况(如细胞色素 c 在细胞凋亡时由线粒体迁移到胞质)。对一些细胞器特别受累的病理情况,亚细胞蛋白质组学研究会更有意义。

三、亚细胞蛋白质组学的技术基础

亚细胞蛋白质组学的技术基础在于亚细胞组分的分离纯化,继以蛋白质组学技术研究。随着亚细胞分级技术、蛋白质组学技术的发展,亚细胞蛋白质组学必将具有越来越深刻的内涵。

(一) 亚细胞组分的分离

在过去的 40～50 年中,细胞的分级分离为生物学家了解细胞器和生物大分子的组成和功能提供了可能。特别是成功地制备了分级的细胞匀浆物(如无细胞体系或体外系统用于研究像 DNA 复制和转录、蛋白质定位和转运、微管的装配以及蛋白质合成的机制等)一系列复杂的机制[23]。进行亚细胞蛋白质组学研究,首先必须完成对亚细胞组分的分离纯化。生物化学和细胞生物学的发展已获得较成熟的亚细胞组分分离技术[24]。

1. 细胞的破碎

亚细胞分级过程包括两个主要的阶段:破碎细胞和细胞成分的分离。破碎培养或组织细胞可以采用渗透压冲击、超声波震荡、机械力研磨或剪切等各种方法来实现。整个过程中可用显微镜监控以确保细胞器结构上的完整。比较温和的破碎程序可以使细胞核、高尔基体、线粒体和其他膜结构保持完整,所得到的匀浆在合适的介质中大部分是有活性的,也就是它保留了其大部分原来的生化活性。然后再根据不同细胞器的不同理化特性自匀浆中分离出各种亚细胞组分。

2. 亚细胞组分的分步分离:差速离心与密度梯度离心联用

匀浆产生的细胞裂解物通常通过一系列差速离心步骤加一个密度梯度离心来分离细胞器和粒子。典型的差速离心顺序在表 19.1 中列出。依次差速离心得到的沉淀再在密

度梯度中进行超速离心。这种逐渐加大离心力的顺序离心及密度梯度的选择将根据细胞成分的大小和密度来实现分离。每种成分与其他成分分离的速度是其大小和形状的函数,称为沉降系数。不同细胞器的大小和沉降特性列于表 19.2[25]。

表 19.1　差速离心形成的沉淀(肝脏)

沉淀	RCF(gav)×时间	内 容 物
P1	1 000g×10min	细胞核,重线粒体,大片细胞膜
P2	3 000g×10min	重线粒体,细胞膜碎片
P3	6 000g×10min	线粒体,溶酶体,过氧化物酶体,完整高尔基体
P4	10 000g×10min	线粒体,溶酶体,过氧化物酶体,高尔基体膜
P5	20 000g×10min	溶酶体,过氧化物酶体,高尔基体膜,大的高密度小泡(如粗面内质网)
P6	100 000g×10min	从内质网而来的所有小泡、细胞膜,高尔基体,核内体等

注:RCF 即相对离心力。

表 19.2　细胞器的大小和沉降特性

亚细胞成分	大小/μm	RCF/g	时间/min
细胞核	4~12	500~1 000	5~10
核膜*		2 000(30 000)	30(5)
线粒体	0.4~2.5	1 000~10 000	10~15
溶酶体	0.4~0.8	6 000~15 000	10~20
过氧化物酶体	0.4~0.8	6 000~15 000	10~20
粗面内质网小泡	0.05~0.35	30 000~100 000	30~60
光面内质网小泡	0.05~0.3	50 000~100 000	30~60
胞膜			
大片	3~20	1 000~3 000	10~15
小泡	0.05~2.0	50 000~100 000	30~60
核内体	0.05~0.4	50 000~100 000	30~60
高尔基体(完整)	1.0~2.0	10 000~20 000	20~30
高尔基体(小泡)	0.05~0.5	50 000~100 000	20~40
肌质网	0.1~1.0	10 000~35 000	20
叶绿体	2~5	1 000~2 000	10
植物线粒体	1~3	5 000~20 000	15

* 核膜的制备物是完整的包膜(鬼影),它的沉降速率取决于制备的方式。

目前有许多种梯度介质成功地用于细胞的分步分离。这些包括蔗糖、Ficoll、Percoll、Nycodenz 和 Metrizoate 等,每种都有其优点和缺点。蔗糖梯度可能是最常用的,这是因为蔗糖便宜、溶解度高,并且能够制备分离绝大多数成分所要求的密度范围的溶液。但是在高浓度时蔗糖黏性大且渗透压高。所有对渗透压敏感的细胞器会在经过不同梯度的过程中反复收缩和吸涨,最后重悬在等渗的介质中。这个效应会损害梯度根据其大小或密度分离细胞器的能力。为了获得较高的密度分布和较小的渗透压和黏度,Nycodenz 和 Percoll 可以作为可接受的替代品。与蔗糖相比,非离子的碘化溶质 Nycodenz(一种三碘苯甲酸衍生物,与 mentrizamide 十分接近)在相同密度下渗透压小得多[26~28]。Nycodenz 在多

数液相介质中都是可溶的，在 pH 2～12.5 的范围内稳定，并且不受二价阳离子的影响。主要缺点是昂贵。Percoll(一种包被了的硅胶)被认为是高浓度蔗糖梯度的另一种有用的替代品。其使用十分简单：细胞被匀浆后，与一种 Percoll 缓冲液混合并在固定角转头中离心。在离心过程中 Percoll 形成自己的梯度。由于它是一种胶体溶液，Percoll 几乎没有渗透压效应；然而 Percoll 粒子沉降得相当快，引起梯度在离心过程中发生保护，这使它在较高速度下使用时比较困难。细胞器通常在低渗透压介质中比在蔗糖溶液中有更低的密度(见表 19.3)。表 19.4 描述了纯化不同细胞器所推荐的梯度介质和离心条件[24]。

表 19.3　亚细胞成分在蔗糖和低渗透压介质中的密度

亚细胞成分	密度/($g \cdot cm^{-3}$)	
	蔗糖	渗低透压介质*
细胞核	>1.30	1.21～1.24
核膜	1.18～1.22	n.a.
线粒体	1.17～1.21	1.15～1.20
Mitoplast 线粒体质	1.44	n.a.
溶酶体	1.19～1.21	1.10～1.15
过氧化物酶体	1.18～1.23	1.19～1.22
粗面内质网	1.18～1.26	n.a.
光面内质网	1.06～1.15	1.03～1.07
细胞膜		
大片	1.14～1.19	1.12～1.15
小泡	1.07～1.16	1.05～1.12
核内体	1.06～1.16	1.05～1.10
高尔基体(完整)	1.05～1.12	1.03～1.08
高尔基体(小泡)	1.05～1.12	1.03～1.08
肌质网	1.04～1.08	n.a.
叶绿体	1.18～1.20	1.10～1.13
植物线粒体	1.16～1.19	n.a.

注：n.a. 未得到数据；

* 这些数据是在 Nycodenz、metrizamid 或 Percoll 中的密度。

差速离心与密度梯度离心联用，不仅广泛用于不同细胞器的分离，业已成功用于一些大分子结构体和多蛋白质复合体的制备。而且，与不同溶解缓冲液系统结合使用，可将一些亚细胞组分进一步细分为更多组分，如 Rout 等[29]从 Trypanosoma brucei 的细胞核里细分出核膜、核孔和核仁组分，Cahill 等[30]分离出大鼠肝脏线粒体的核糖体。

3. 免疫共沉淀法(Co-IP)分离信号复合体(signaling complex)

细胞信号传导一直是细胞分子生物学的一大热门课题。传统的信号传导研究对单一信号分子上下游蛋白质和蛋白质之间作用已形成一套模式，对许多信号通路的发现做出重要贡献。随着信号通路之间 cross-talk 的发现，信号通路的复杂性已超出初期的想像，细胞内信号网络的概念已被普遍接受。对细胞内信号传导的研究必将要求新的思维模式和技术要求。蛋白质组学由于其系统性、整体性思维模式，已被迅速引入细胞信号传导研

究。信号复合体不能通过密度梯度离心纯化，而免疫沉淀（immunoprecipitation）正适应了这一要求。将目标信号分子蛋白从细胞或亚细胞组分中沉淀后，由于蛋白质和蛋白质之间相互作用，所有共沉淀下的其他蛋白质可假定为参与信号复合体，经蛋白质组学技术分离鉴定后，可进一步经细胞分子生物学实验验证其在信号传导中的功能[16~18]。

表 19.4　纯化不同细胞器所推荐的梯度和离心条件

成分	沉淀[a]	RCF(gav)×时间	梯度[b]	污染物	参考文献
细胞核	P1	100 000g×60min	60%蔗糖屏障	无	Blobel and Potter 1996[40]
	P1/匀浆物	15 000g×60min	20%～50% dc Nycodenz	无	Comerford *et al*. 1985[41]
大片细胞膜	P1	160 000g×3h	37%/60%蔗糖	线粒体	Graham *et al*. 1990[42]
线粒体	P2	3 000g×10min	无	细胞核	
	P2+P3+P4	50 000g×2h	20%～40% dc Nycodenz	过氧化物酶体	Graham *et al*. 1990[42]
	匀浆物	37 000g×30s	20%～52% dc Percoll	过氧化物酶体	Graham 1984[43]
溶酶体	P4+P5	50 000g×2h	10%～50% dc Nycodenz	细胞膜	Graham *et al*. 1990[42]
	P4+P5	37 000g×30min	15%～85% c Percoll	线粒体	Graham 1984[43]
过氧化物酶体	P4+P5	95 000g×2h	34%～47% dc Nycodenz	线粒体	Wattiaux and Wattiaux-De Coninck 1983
高尔基体	P4+P5	160 000g×1h	10%～44% dc 蔗糖	细胞膜小泡/光面内质网	Graham 1984[43]
	P4+P5	50 000g×2h	10%～50% dc Nycodenz	溶酶体	Graham *et al*. 1990[42]
光面内质网/粗面内质网	P6	150 000g×1h	20%/45%蔗糖加15mmol/L 氯化铯	光面内质网/核内体	Bergstrand and Dallner 1986 [44]
细胞膜小泡	P4+P5/P6	100 000g×4h	10% ～30% c metrizamide	光面内质网/核内体	Graham and Winterbourne 1988[45]
核内体	HOM/P6	85 000g×45min	5%～35% dc Nycodenz	光面内质网/细胞膜小泡	Kindberg *et al*. 1984[46]
肌质网	P2	120 000g×2h	20%～38% dc 蔗糖	线粒体	Meissner 1984[47]
核膜	P1	100 000g×1h	10%～55% dc 蔗糖	无	Key *et al*. 1972[48] Graham *et al*. 1990[42]
线粒体外膜	P2	115 000g×1h	23%～43% dc 蔗糖	线粒体内膜	Coty *et al*. 1979[49]

a. P1～P6 同表 19.1； b.(dc)：非连续；(c)：连续。

4. 亲和层析法分离不同亲和特性的亚细胞组分

亲和层析法是一类经典的蛋白质分离纯化技术。选择一定的基质与配基,可将细胞或亚细胞内不同亲和特异性的蛋白质群分离出来并继以蛋白质组学研究。Lopez 等[19]通过 Ca^{2+} 螯合层析柱、ConA 亲和层析柱及苯基琼脂糖凝胶树脂分别分离出大鼠肝组织线粒体中的钙结合蛋白质、糖蛋白和疏水性膜蛋白,并进行了蛋白质组学研究。需特别说明的是,细胞或亚细胞组分中不同种类的翻译后修饰的特性可望用来分离各类被修饰的蛋白质。

5. 分离纯化效果评价

亚细胞分离纯化是否成功,可以用光学或电子显微镜监测,或者检测蛋白质浓度或合适的标志酶。如 Murayamad 等人[31]就通过 Western 印迹检测不同亚细胞成分的标志性蛋白质或酶,鉴别了由 Nycodenz 密度梯度离心一步法分离到的不同亚细胞组分。在匀浆和纯化过程中跟踪合适的标志酶的活性可以确定目的细胞器的产率和污染其他细胞器成分的程度。而且,在整个纯化过程中确定蛋白质的浓度,在每一步计算比活性(活性与蛋白质的比率),可以估计富集程度或纯化倍数。主要细胞成分的标志酶列于表 19.5[25]。

表 19.5 哺乳动物细胞膜的主要酶特点

膜成分	酶	说明
细胞膜		
顶端膜	5′-核苷酸酶,亮氨酸氨肽酶,碱性磷酸酶	在一些细胞中该膜可能与一些组织特异性功能有关,如在肠道细胞中有糖转运功能
基底膜	Na^+/K^+-ATP 酶,激素控制的腺苷酸环化酶	肝组织中的 Na^+/K^+-ATP 酶主要存在于毗邻的膜上,环化酶在血窦膜上
通用	5′-核苷酸酶,Na^+/K^+-ATP 酶	这些最常使用的通用标志在组织培养的细胞膜上并不总是容易分析
线粒体	琥珀酸脱氢酶	可以用细胞色素 c 或 P-碘硝基四唑紫作为电子受体检测;在培养细胞中的水平倾向于较低
溶酶体	β-半乳糖苷酶,酸性磷酸化酶,芳基硫酸酯酶	
过氧化物酶体	过氧化氢酶,尿酸酶	可以用 *N*-乙酰葡糖胺或卵清蛋白作为受体来检测
高尔基体	半乳糖苷转移酶	
内质网	NADPH(或 NADH)-细胞色素 c 还原酶	只有肝和肾有较高水平的葡萄糖-6-磷酸酶是否存在内质网污染是很难估计的,因为膜有一定的连续性
核膜	核苷酸三磷酸酶	
线粒体外膜	单胺氧化酶	
线粒体内膜	细胞色素氧化酶	

（二）蛋白质组学技术

蛋白质组学技术主要在于蛋白质分离技术与蛋白质鉴定技术的结合。不断改进的蛋白质分离技术和不断更新换代的质谱技术以及快速发展的生物信息学技术，使得当今的蛋白质组学研究日新月异。这些技术的有机结合可适应亚细胞蛋白质组学研究的不同目的和不同特性的要求。

传统的 2-DE-MS 技术中，2-DE 的一些不足与发展前景尽管目前虽有争议[13]，但它仍有许多不可忽视的优越性。如可提供蛋白质的 pI 和 M_r 参考数值，有助于蛋白质的鉴定[32]。而且 2-D 胶上常见的蛋白质同形异构体（isoform）多是蛋白质的翻译后修饰结果[33,34]。

目前，飞速发展的质谱技术对酶解肽段的检测和检索功能的加强，对蛋白质量及纯度的要求降低。许多研究者在分析蛋白质种类较少时（如免疫共沉淀蛋白质复合物），通过单向凝胶电泳（one-dimensional gel electrophoresis, 1-DE）或液相色谱（liquid chromatography, LC）或二者结合分离蛋白质后直接进行蛋白质混合物的质谱鉴定，而不再依赖于 2-DE[35-37]。

对亚细胞蛋白质组学研究有着重大意义的蛋白质组学技术是多维色谱与串联质谱联用的 2D-LC-MS/MS 技术用于直接分析大蛋白质复合体（direct analysis of large protein complexes, DALPC）[13,38]。蛋白质复合物酶解后的肽段混合物经强阳离子交换柱和反相柱两阶段分离后，由电喷雾离子化串联质谱（ESI-MS/MS）分析获得大量肽段的 MS/MS 谱图，经数据库检索后可一次性鉴定出数百种甚至更多的蛋白质。2D-LC-MS/MS 技术避免了双向凝胶电泳技术分离蛋白质过程中的等电点极限、疏水性、相对分子质量范围等限制性因素，且可检测 10 000∶1 丰度范围内的低丰度肽段[38]。采用 2D-LC-MS/MS 技术直接分析亚细胞组分蛋白质混合物，将是亚细胞蛋白质组学研究的重要发展方向。

（三）蛋白质亚细胞定位实验技术

亚细胞分离技术本身已提供了蛋白质的亚细胞定位信息，毋庸怀疑所鉴定出的蛋白质绝大部分定位在相应的亚细胞组分中或与该组分相关。但由于完全纯化的亚细胞组分很难获得也很难证明，为了验证亚细胞蛋白质组学的研究结果，如新发现蛋白质的亚细胞定位及其在亚细胞结构中的可能功能，需要细胞或分子生物学的补充实验予以证实。如应用 Western 印迹验证所发现蛋白质在亚细胞组分中的存在；TAP（tandem affity purification）-tag 方法可望解决免疫共沉淀中的非特异性结合问题[39]。蛋白质精确亚细胞定位技术除亚细胞组分分离外还有电镜技术和荧光显微镜技术[40]。可采用免疫电镜（immunoelectron microscopy）[21,22]或免疫荧光（immunofluorescence）技术确认可疑蛋白质在细胞内的定位；甚至可以通过表达 myc-tag 或绿色荧光蛋白（GFP）融合蛋白，分析它们在细胞中的定位[20,41]。基因缺失突变不但可以验证蛋白质的亚细胞的定位，还可以就该蛋白质的生物功能提供证据[13]。

（四）蛋白质亚细胞定位预测技术

虽然以上这些技术可提供蛋白质的相对精确的亚细胞定位信息，但这些实验技术在

目前多是耗时的，且主观性强，可重复性差[40]。在功能基因组学尤其是蛋白质组学研究时代，随着大量新蛋白质的发现，单纯依靠这些实验技术验证蛋白质的亚细胞定位是不现实的。近 10 年来，许多学者尤其是一些生物信息学家根据蛋白质序列信息在蛋白质亚细胞定位的预测方面进行了广泛研究并取得了一系列有意义的结果[42~45]。表 19.6 列举了互联网上可用的一些主要亚细胞定位预测软件或服务器。

表 19.6　互联网上常用的亚细胞定位预测软件

软件名称	网址(URL)	预测功能	预测依据	参考文献
ChlorpP	http://www.cbs.dtu.dk/services/Chlorop/	叶绿体蛋白质	识别 CTP	[46]
MitoProt	http://www.mips.biochem.mpg.de/cgi-bin/proj/medgen/mitofilter	线粒体蛋白质	识别 MTP	[47,48]
SignalP	http://www.cbs.dtu.dk/services/SignalP/	分泌蛋白	识别 SP	[49]
predictNLS	http://maple.bioc.columbia.edu/predictNLS/	核蛋白质	识别 NLS	[50]
Predotar	http://www.inra.fr/Internet/Produits/Predotar/	线粒体和叶绿体蛋白质	识别 MTP 和 CTP	[51]
TMHMM	http://www.cbs.dtu.dk/services/TMHMM-2.0/	膜蛋白质	识别跨膜螺旋和拓扑结构	[52]
HMMTOP	http://www.enzim.hu/hmmtop/	膜蛋白质	识别跨膜螺旋和拓扑结构	[53]
SobLoc	http://www.bioinfo.tsinghua.edu.cn/SubLoc	12 种亚细胞定位	氨基酸组成分析	[54]
NNPSL	http://predict.sanger.ac.uk/nnpsl	5 种亚细胞定位	氨基酸组成分析	[55]
TargetP	http://www.cbs.dtu.dk/services/TargetP/	4 类亚细胞定位	SignalP、ChloroP 和 MTP 识别三者相结合	[56]
PSORT	http://psort.nibb.ac.jp/	12 种亚细胞定位	氨基酸组成与分拣信号识别相结合	[57~61]

注：CTP：叶绿体运输肽(chloroplast transit peptide)；NLS：核定位信号(nuclear localization signals)；MTP：线粒体引导肽(mitochondrial targeting peptide)；SP：信号肽(signal peptide)。

一般而言，各种亚细胞定位预测方法都包括四个步骤：首先，建立客观、有代表性的各种已知亚细胞定位的训练用蛋白质数据系列(training dataset)；其次，从这些蛋白质数据系列中抽提出各种特征信息参数；再次，根据特征信息参数采用一些算法建立算法模型以计算比较可疑蛋白质某种亚细胞定位的可能性；最后，采用检验用蛋白质数据系列(testing dataset)结合不同评价方法对所建立算法模型的预测结果进行评价[45]。

迄今为止，亚细胞定位预测主要基于蛋白质氨基酸序列的三个方面的特性或它们之间的相互结合：氨基酸组成特点、分拣信号的识别以及氨基酸序列的进化保守性[43]。许

多研究表明各种细胞器有偏好性的氨基酸组成[62]。针对这一特性,不同学者又采用不同方式进行蛋白质的亚细胞定位预测,如 Cedano 等人采用了判别式分析方式(discriminant analysis)[63]; Reinhardet 和 Hubbard 采用了人工神经网络方式(artificial neural network)[55]; Chou 和 Elrod 采用了一种协变式判别式运算法则(covariant discriminant algorithm)[64~66];而 Yuan 等人采用了一种 Markov 链模型(markov chain Model)[67]。蛋白质在胞浆合成后,须在 N 端分拣信号(sorting signal)的指引下到达不同亚细胞部位发挥生物学功能。各种分拣信号如信号肽(signal peptide)、线粒体引导肽(mitochondrial targeting peptide)、叶绿体运输肽(chloroplast transit peptide)和核定位信号(nuclear localization signals)等已被广泛单独或联合用于蛋白质的亚细胞定位预测[44]。如 ChloroP 通过识别叶绿体运输肽预测蛋白质的叶绿体定位[46];MitoProt 通过识别线粒体引导肽预测线粒体或非线粒体定位[47,48];SignalP 通过识别信号肽用于预测分泌蛋白质甚至预测信号肽酶切位点[49]。而 TargetP 通过 SignalP、ChloroP 和线粒体引导肽识别三者相结合将蛋白质亚细胞定位预测为分泌蛋白质、线粒体蛋白质、叶绿体蛋白质和其他四类[56]。不断完善的 PSORT 结合氨基酸组成分析与分拣信号识别预测包括各种细胞器的 12 种亚细胞定位方式[57~61]。Stanford 等人[68]认为,真核生物同工酶的不同亚细胞定位与同工酶之间的氨基酸序列差异有关,将这些差异序列进一步与真菌或原核生物的同源蛋白质序列比较,将会明确潜在的定位信号区域。据此,他们收集了真核蛋白质定位附加区域(additional domains for eukaryotic protein targeting,ADEPTS)。Marcotte 等人[69]根据同源蛋白质在系统发生图谱上亚细胞定位分布情况,将线粒体蛋白质分为原核生物源性、真核生物源性及物种特异性三类。

蛋白质亚细胞定位预测已取得一定的成绩,如大多软件对检验用蛋白质数据系列的预测成功率超过 65%,TagertP 对植物和非植物蛋白质的预测成功率更是分别高达 85% 和 90%。利用 TargetP 预测出约 10%的拟南芥和人基因产物为分泌蛋白质[56]。尽管如此,生物信息学家对目前的预测软件能否普遍应用仍持怀疑态度,要求对预测的结果慎重对待[45]。这主要是因为目前的亚细胞定位预测技术还存在许多问题,主要如蛋白质数据库中亚细胞定位信息不完善导致训练用数据代表性不强;氨基酸组成信息难以有效提取;蛋白质运输与分拣机制尚未完全清楚,各类分拣信号尚缺乏足够代表性;新蛋白质氨基酸序列的 N 端难以确定从而妨碍分拣信号的识别,等等。蛋白质的亚细胞定位预测技术仍需不断发展完善。

亚细胞蛋白质组学研究与蛋白质亚细胞定位预测技术相结合可以相互验证,相互促进。一方面,亚细胞蛋白质组学研究获得的大量亚细胞组分蛋白质,难以全部经实验技术确证其亚细胞定位;经过系统性的亚细胞定位预测,既可以检测亚细胞组分中整体蛋白质的亚细胞定位特性,又可根据单个蛋白质亚细胞定位预测结果与整体亚细胞组分特性的符合情况,选择出有意义蛋白质进一步研究。另一方面,各种亚细胞组分的蛋白质组学研究结果将会富集特异的亚细胞组分蛋白质,对它们的预测结果将反映预测软件的实际预测效果和对各类不同亚细胞定位的蛋白质的预测效果;而且高质量的亚细胞蛋白质组学研究结果提供的蛋白质的亚细胞定位信息可进一步用作训练数据,从而完善预测技术。

四、亚细胞蛋白质组学研究进展

世界范围内的亚细胞蛋白质组学研究方兴未艾。不同生物体的不同器官、组织及多种细胞株的不同亚细胞组分的蛋白质组学研究时有报道。表 19.7 列举了目前互联网上的部分亚细胞组分的蛋白质组学数据库网址。

表 19.7　互联网上有关亚细胞组分的蛋白质组学数据库

亚细胞组分名称	蛋白质组学数据库网址	参考文献
人肝细胞核	http://us.expasy.org/cgi-bin/ch2d-compute-map? NUCLEI _ LIVER _ HUMAN	[70]
人胎盘溶酶体	http://www.health.adelaide.edu.au/2Ddatabase/front2.htm	[71]
人胎盘线粒体	http://www.dsv.cea.fr/thema/MitoPick/Mito2D.html	[72]
拟南芥线粒体	http://www.gartenbau.uni-hannover.de/genetik/AMPP	[73]
线粒体	http://www.mips.biochem.mpg.de/proj/medgen/mitop/	
Hela 细胞核仁	http://us.expasy.org/cgi-bin/map2/def? NUCLEOLI _ HELA _ 2D _ HUMAN	[74]
Hela 40S hnRNP	http://proteomics.cancer.dk/2Dgallery/40S _ hnRNP.html	
人蛋白酶体	http://proteomics.cancer.dk/cgi-bin/gel _ launcher? Proteasomes + (IEF) £ proteasomes _ ief	
人核糖体	http://proteomics.cancer.dk/cgi-bin/gel _ launcher? Ribosomes + (NEPHGE) £ ribosomes	
小鼠线粒体	http://proteomics.cancer.dk/cgi-bin/gel _ launcher? Mouse + mitochondria + (IEF) £ mouse _ mitochondria	
小鼠肝细胞核	http://us.expasy.org/cgi-bin/ch2d-compute-map? NUCLEI _ LIVER _ MOUSE	[75]

1. 细胞器的蛋白质组学研究

几乎各种类型的细胞器都有蛋白质组学研究报道。如 Rabilloud 等人[72]分离出人胎盘线粒体,鉴定出 53 种蛋白质;Chataway 等人[71]从胎盘溶酶体中鉴定出 19 种蛋白质,并将 3 种未知蛋白质进行了 N 端测序;Muller 等人[76]从 Burkitt 淋巴瘤 BL60 细胞株分离出细胞核,鉴定了 33 种蛋白质并获得 3 种未知蛋白质的氨基酸序列信息;Fialka 等人[20]从 MDCK Ⅱ细胞的内吞小泡中发现 8 种已知蛋白质和 9 种未知蛋白质,并鉴定出其中一种未知蛋白质。关于线粒体和高尔基体的蛋白质组学研究将在相关章节中详细介绍。

2. 大分子结构体和多蛋白质复合体的蛋白质组学研究

大分子结构体和多蛋白质复合体作为细胞内的一些基本结构与功能单位,对其蛋白质组成的了解具有重要意义。对大分子结构体和多蛋白质复合体的蛋白质组学研究已有众多报道。如 Gerner 等人[77~79]从人的白细胞、羊膜细胞和肝组织细胞三类细胞的核基质中共鉴定出了 15 种核基质蛋白;Neubauer 等人[41]从 HeLa 细胞中分离出剪接体(splicesome),鉴定出 17 种已知剪接体蛋白质,新发现 4 种蛋白质与剪接体相关,并通过

EST 库搜索发现 19 种与剪接体相关的新蛋白质；Heller 等人[17]从 T 细胞抗原受体(T-cell antigen receptor ,TCR)信号复合体中鉴定出 7 种新的蛋白质成分，可望用于药理学研究。对核孔复合体和核仁的蛋白质组学研究参见相关章节。

3. 酵母的亚细胞蛋白质组学研究

酵母作为一种被广泛应用的基因组信息较为完善的模式生物，对其亚细胞蛋白质组学研究已取得重要进展。如 Wigge 等人[22]鉴定了酵母纺锤体组分中的 23 种蛋白质，其中 11 种新基因产物中的 8 种共定位于纺锤体；Link 等人[80]对酵母 80S 核糖体大亚基的蛋白质组学研究，鉴定出已知的 74 种蛋白质，并鉴定了一种新的核糖体蛋白质；Rout 等人[21]对酵母核孔复合体的蛋白质组学研究发现了 40 种与核孔复合体相关的蛋白质。Yoon 等人[81]通过串联亲和纯化技术(tandem affinity purification, TAP)纯化出酿酒和裂殖两类酵母的 anaphase-promoting complex (APC)，均鉴定出 13 种蛋白质。裂殖酵母的 Cdc5p 与其酿酒酵母的类似物 Cef1p 是 Myb 相关的蛋白质，与 mRNA 前体的剪接有关。Ohi 等人[82]证明了它们存在于稳定的多蛋白质复合体中，并且通过 TAP 技术纯化出 Cdc5p 与 Cef1p 复合体，继以多蛋白质复合体直接分析技术(direct analysis of large protein complex, DALPC)分析后，发现二者具有几乎相同的蛋白质组成，主要包括 Sm、U2、U5 snRNP 的蛋白质成分，而且发现含有 U2、U5、U6 snRNAs。另外，还鉴定出一些新蛋白质与 mRNA 前体的剪接有关。

4. 结语

随着各种亚细胞组分分离技术和蛋白质组学技术的发展，各种新的亚细胞组分的蛋白质组学研究的报道层出不穷；而且，同一亚细胞蛋白质组学的结果也不断被新的研究所更新完善。例如，对大鼠肝组织线粒体的蛋白质组学研究，在 3 年内，美国、日本和瑞士的 3 个研究小组分别用不同方法从不同角度逐步完善了对大鼠肝组织线粒体的蛋白质组学研究[19,31,83]。

五、总　　结

亚细胞蛋白质组学通过亚细胞分离技术、蛋白质组学技术以及相关的分子、细胞生物学的验证性技术对各种亚细胞组分进行蛋白质组学分析。

随着各种分离技术的发展，亚细胞蛋白质组学具有越来越深刻的内涵。除了传统的细胞器的蛋白质组学研究，对各种大分子结构体以及多蛋白质复合体(包括信号复合体)的蛋白质组学研究也取得了重大进展。

亚细胞蛋白质组学的研究成果证明了亚细胞蛋白质组学研究的重要意义。表现在：可以富集低丰度蛋白质从而完善全细胞蛋白质组；可以提供蛋白质的亚细胞定位信息；可以加深对亚细胞组分的结构、功能的理解；可以促进对细胞生理分子机制的理解；通过亚细胞蛋白质组学差异图谱比较分析对疾病病理机制等的研究具有特别意义。

结合亚细胞分离技术，蛋白质组学技术及其系统性、整体性思维方式，亚细胞蛋白质组学必将在细胞生物学、分子生物学领域发挥重要作用。亚细胞蛋白质组学正是蛋白质

组学通向细胞生物学、分子生物学的桥梁,将促进生命科学研究进一步走向深入。

第二节　细胞膜的蛋白质组学研究进展

生物膜系统包括细胞质膜和真核细胞中的细胞内膜,具有重要生物功能和特征性的分子结构特点。由于膜蛋白质的高度疏水性、难溶解性,对细胞膜的蛋白质组学研究反映了蛋白质组学研究面临的一些技术挑战。

一、细胞膜的结构与功能

细胞膜(cell membrane)又称质膜(plasma membrane),是细胞的重要组分,包围在所有细胞的表面,为一层极薄的膜。在真核细胞中,除了质膜以外,细胞内还有构成各种细胞器的膜,称为细胞内膜。细胞质膜与细胞内膜统称为生物膜(biomembrane)。

质膜最基本的作用是维持细胞内微环境的相对稳定,并与外界环境不断地进行物质交换、能量和信息的传递,对细胞的生存、生长、分裂分化都至关重要。细胞内膜更是各种细胞器发挥生物功能的基础。生物膜与生命科学中许多基本问题以及有关的一些迫切需要解决的实际问题都有密切关系。例如,细胞起源、细胞分裂分化、细胞识别、免疫、物质运输、信息传递、代谢调控、能量转换、神经传导以及肿瘤发生等无不与膜有关[84,85]。

不同的细胞质膜与不同的内膜各有其不同的功能,但它们的化学组成和分子结构却有着共同的特征。细胞膜主要是由脂类和蛋白质(包括酶)组成。所有的膜都包含上述两类成分。在质膜的表面还普遍存在有少量的糖类,形成糖脂和糖蛋白。大体上说,脂类约占总量的 50%,蛋白质约占 40%,糖类占 1%～10%。实际比例随着膜种类不同而有很大差别。膜中脂类和蛋白质含量的变化与膜的功能有关。一般而言,膜所含蛋白质越多,膜的功能越复杂。相反,膜的功能越简单,所含蛋白质的种类和数量越少。随着对膜研究的深入,对膜的分子结构已先后提出“蛋白质-脂类-蛋白质”三明治式结构模型、单位膜模型和流体镶嵌模型等[84,85]。

二、膜蛋白质的特点

根据膜蛋白分离的难易及其与脂分子的结合方式,膜蛋白可分为两大基本类型:外在膜蛋白(extrinsic membrane proteins)或称外周膜蛋白(peripheral membrane proteins)和内在膜蛋白(intrinsic membrane proteins)或称整合膜蛋白(integral membrane proteins)。外在膜蛋白为水溶性蛋白质,靠离子键或其他较弱的键与膜表面的蛋白质分子或脂分子结合。因此,只要改变溶液的离子强度甚至提高温度就可以从膜上分离下来,膜结构并不被破坏。内在膜蛋白以不同程度嵌入脂双层的内部,多为横跨全膜,其疏水区域与脂双层中脂类分子的疏水尾部相互作用,亲水区域暴露在膜的一侧或两侧表面。因此,这些蛋白质也具有双型性,其中疏水性氨基酸比例较高。在内在膜蛋白的跨膜疏水部分几乎都含有 1 个或 1 个以上的 α-螺旋结构片段。内在膜蛋白与膜结合非常紧密,不易分离,只有用去垢剂(detergent)使膜崩解后才可分离出来,但分离后易失去其正常的构型。因此,膜蛋

白，特别是内在膜蛋白的分离提纯远远少于外在蛋白和非膜蛋白的研究[84,85]。

三、膜蛋白质的双向凝胶电泳技术研究

膜的重要生物学功能和结构特点，已激发了许多学者对膜的蛋白质组学研究。但膜蛋白质的不易分离和疏水特性对传统的蛋白质组学技术尤其是双向凝胶电泳技术提出了严峻挑战。双向凝胶电泳技术作为蛋白质组学研究中最重要、最常用的蛋白质分离技术，对膜蛋白的分离能力却很差，膜蛋白甚至在双向凝胶电泳图谱上缺失[86,87]。这主要可能是由于膜蛋白的难溶解性[88]以及与固化 pH 梯度（immobilized pH gradients, IPG）胶条基质间的疏水作用造成[89,90]。

为了增强双向凝胶电泳系统溶解膜蛋白的能力，各种新的破膜剂（chaotropes）、表面活性剂（surfactants）、还原剂（reducing agents）不断被应用于膜蛋白的样品制备[91]。如新的破膜剂硫脲（thiourea）与尿素（urea）的联用，在双向凝胶电泳图谱上分离到单一应用尿素时未发现的一些核蛋白、内在膜蛋白和微管蛋白等[92]。β 硫代型两性离子表面活性剂（sulfobetaine-type zwitterionic surfactants）如 Chaps 已被证明较非离子变性剂 NP-40 或 Triton X-100 等对双向凝胶电泳分离蛋白质有更好的效果[93~95]。使用 TBP（tributyl phosphine）代替常规还原剂二巯基乙醇（dithiothreitol, DTT）应用于双向凝胶电泳，增强了中国仓鼠卵巢细胞蛋白质或羊毛角蛋白的溶解分离效果[96,97]，在双向凝胶电泳图谱上分离到四种相对分子质量和 pI 值相近的角蛋白。各种新试剂的引用及不同试剂相互组合的应用已使双向凝胶电泳技术溶解分离膜蛋白的能力得到一定发展。

除了上述新试剂的应用外，针对膜蛋白组分的特点，一些新策略已被发展应用于膜蛋白的双向凝胶电泳分析研究，如膜蛋白组分的分级抽提[98]、采用双向变性剂 Triton X-144[99]或碱性 Na_2CO_3[100] 预处理膜蛋白组分分离内在膜蛋白技术以及采用有机溶剂抽提膜蛋白[101]等。Molloy 等[98]首先用 Tris-base 抽提出大肠杆菌的易溶蛋白质，接着用常规裂解溶液（8mol/L 尿素，4% Chaps，100mmol/L DTT，0.5% CA）抽提出一般疏水性蛋白质，其次用含硫脲、TBP 和 sulfobetaines 等试剂的强裂解液抽提难溶蛋白质，剩下不溶组分用含强表面活性剂 ASB14 的裂解液（7mol/L 尿素，2mol/L 硫脲，1% ASB14，2mmol/L TBP，40mmol/L Tris，0.5% CA）溶解。所得各个组分分别经双向凝胶电泳分析后，在后两个组分中新富集鉴定出一批含 OmpA 的外膜内在膜蛋白。Molloy 等[102]接着采用碱性 Na_2CO_3 预处理技术分离出大肠杆菌膜内在膜蛋白，用强裂解液（7mol/L 尿素，2mol/L 硫脲，1% ASB14，2mmol/L TBP，40mmol/L Tris，0.5% CA）溶解后经双向凝胶电泳技术（pH 4～7，分子质量 10 000～80 000Da）分离后，鉴定出 21 种内在膜蛋白，约占 SwissProt 蛋白质库中预测的大肠杆菌内在膜蛋白的 80%。Santoni 等[103,104]以模式生物拟南芥叶膜蛋白质为研究对象，系统比较优化了各种应用于双向凝胶电泳分析的膜蛋白质的抽提和溶解方案，发现碱性 Na_2CO_3 预处理膜蛋白组分分离出的内在膜蛋白经含有强表面活性剂 C8ϕ 的裂解液溶解后，用免疫学方法鉴定出如水通道蛋白、纤维素酶、ATPase 等内在膜蛋白[103]，并表明膜蛋白组分中脂质含量影响变性剂对膜蛋白的溶解效果，需进行预处理[104]。

四、膜蛋白质的非双向凝胶电泳技术研究

由于传统双向凝胶电泳技术分离膜蛋白质的局限性，愈来愈多的学者尝试采用一些非双向凝胶电泳技术研究膜蛋白质。随着质谱对蛋白质或肽段混合物分析鉴定技术的发展，单向 SDS-PAGE 直接与质谱技术结合分离鉴定膜蛋白质混合物，已成为一种趋势。而且，单向 SDS-PAGE 直接分离膜蛋白质避免了采用 IPG 进行等电聚焦过程造成的膜蛋白质的丢失。Galeva 和 Altermann[105]采用单向 SDS-PAGE 与质谱技术结合分离鉴定大鼠微体(microsome)"鬼影"(ghost)结构中的膜蛋白质。考马斯亮蓝染色的蛋白质条带经胶内酶解后由 MALDI-MS 共鉴定出已知 34 种微体膜蛋白质中的 22 种，而同时进行的双向凝胶电泳技术只分离鉴定出 3 种微体膜蛋白质。

另外，用于直接分析大蛋白质复合体(direct analysis of large protein complexes, DALPC)的多维色谱与串联质谱联用的 2D-LC-MS/MS 技术[13,38]，避免了双向凝胶电泳技术分离蛋白质过程中的等电点极限、疏水性、相对分子质量范围等限制性因素。Washburn 等[106]利用这一技术鉴定出的 1 484 种酵母蛋白质中包括 131 种整合膜蛋白，经预测均含有三个或以上的跨膜区结构域(transmembrane domain)。这一技术必将广泛应用于膜的蛋白质组学研究。

五、总　　结

虽然膜蛋白的疏水性、难溶性限制了传统双向凝胶电泳技术对膜蛋白质的分离，但是各种特性变性剂的发现、应用以及膜蛋白质的不同抽提方案的实施正逐步改善这种困境，并由此促进了蛋白质组学研究中蛋白质分离技术的发展。各种非双向凝胶电泳技术的发展将使膜的蛋白质组学研究得到重大突破，同时也反映了蛋白质组学技术的发展趋势。

第三节　高尔基体蛋白质组学研究进展

高尔基体是真核细胞内的一种重要细胞器，是细胞内大分子运输的一个主要的交通枢纽，主要参与细胞分泌活动，是糖类合成和蛋白质翻译后修饰的重要场所。对其蛋白质组成的了解将促进对高尔基体结构和功能的深入理解。以 Howell 研究小组为主，已对高尔基体进行了一系列的蛋白质组学研究。

一、高尔基体的结构和功能

高尔基体(Golgi body)又称高尔基器(Golgi apparatus)或高尔基复合体(Golgi complex)，是普遍存在于真核细胞内的一种细胞器[107,108]。它不仅存在于动植物细胞中，而且也存在于原生动物和真菌细胞内。在光镜下可见脊椎动物大多数细胞的高尔基体呈复杂的网状结构，少数细胞如卵细胞、精细胞以及大多数无脊椎动物细胞和植物细胞具分散的高尔基体。据估计，每个肝细胞约有 50 个高尔基体，约占细胞质体积的 2%，植物细胞

有数十个到数百个分散在整个细胞质中。高尔基体的形态、大小、分布在不同细胞甚至细胞生长的不同阶段都有很大不同,但在已成熟的细胞却显示有共同的形态特征。

电子显微镜观察到的高尔基体最富有特征的结构是由一些(常常 4~8 个)排列较为整齐的扁平膜囊(saccule)堆叠在一起,构成了高尔基体的主体结构,扁囊多呈弓形,也有的呈半球形或球形,均由光滑的膜围绕而成,在其周围常结合有一些小管和大小不同的有被或无被的囊泡。高尔基体是一种有极性的细胞器,这不仅表现在它在细胞中往往有比较恒定的位置与方向,而且物质从高尔基体的一侧进入,从另一侧输出。在很多细胞中,高尔基体靠近细胞核的一面,扁囊弯曲成凸面又称形成面(forming face)或顺面(cisface),面向细胞质膜的一面常呈凹面又称成熟面(mature face)或反面(trans face)。顺面多与 ER 膜相联系,接近于顺面的扁囊膜在形态与染色性质上与 ER 相似;在反面的扁囊的形态与化学组成很像质膜。已证实高尔基体是由许多功能不同的区隔(或称房室)(compartment)所组成的完整体系,至少可分为高尔基体顺面的网状结构(cis-Golgi network, CGN)、中间膜囊(medial Golgi)和反面的管网结构(trans Golgi network, TGN)。Rothman 等[109]应用单克隆抗体免疫细胞化学电镜技术证明每个区隔内含有的酶不同,功能不同,如 *N*-乙酰葡萄糖胺转移酶只在中间膜囊中,而半乳糖基转移酶只存在于反面的扁囊。

高尔基体的主要功能[107,108,110]是将内质网合成的多种蛋白质进行加工、分类与包装,然后分门别类地运送到细胞特定的部位或分泌到细胞外。内质网上合成的脂类一部分也要通过高尔基体向细胞质膜和溶酶体膜等部位运输,因此,可以认为,高尔基体是细胞内大分子运输的一个主要的交通枢纽。此外,高尔基体还是细胞内糖类生物合成的主要场所,如完成糖蛋白的合成、多糖的合成,以及氨基多糖的硫酸化,都在高尔基体内进行。

二、高尔基体的蛋白质组学研究

Taylor 等[111]根据高尔基体所占细胞容积比估计,高尔基体至少有 1 000 种组成性蛋白质。而根据蛋白质的功能活性或抗原特性从多种哺乳动物的各种组织中只鉴定出不到 200 余种高尔基体蛋白质。全面了解同一物种组织高尔基体的蛋白质组成必将促进深入理解高尔基体的特征结构和所有已知或未知功能。近 5 年来对某些物种组织的高尔基体的蛋白质组学研究已有多篇报道。

Howell 研究小组对大鼠肝组织高尔基体蛋白质组学进行了一系列研究。由于高尔基体是细胞内合成蛋白质运输的一个主要的交通枢纽,在生理稳定状态下,大鼠肝组织高尔基体中约 50% 的蛋白质属于一过性运输蛋白质[112]。1977 年,Taylor 等[113]采用放线(菌)酮(cycloheximide)预处理 SD 大鼠抑制蛋白质合成,待高尔基体中一过性蛋白质转运至各自目的地后,结合差速离心和蔗糖密度梯度离心获得高度富集的高尔基体组分(stacked Golgi fraction, SGF1)。通过高尔基体和其他细胞器一些标志性蛋白质或酶的定量检测证明对照组大鼠肝组织高尔基体组分(CTL SGF1)较差速离心中的去核上清组分(postnuclear supernatant, PNS)富集了 300~400 倍,而放线(菌)酮处理组获得的 CHX SGF1 较 PNS 富集了 400~700 倍。继之以双向凝胶电泳(2-DE)和银染技术获得 CTL PNS, CTL SⅡ(分离过程中的中间组分), CTL SGF, CHX SGF, CHX SGF-high pH 上清组

分和 CHX SGF-high pH pellet 6 个组分的 2-DE 图谱[114]。通过分析各图谱上蛋白质点的匹配情况和富集程度将其中丰度较高的 173 个蛋白质点分为 3 类:80 个运输性蛋白质(cargo protein),44 个与高尔基体相关的胞质蛋白质(cytosolic Golgi-associated protein)和 49 个高尔基体居留蛋白质(resident Golgi protein),并利用强碱液进一步将高尔基体居留蛋白质中的胞质面和腔面的周边蛋白质与膜蛋白质抽提区分出来。为了分离鉴定出更多的高尔基体中的低丰度蛋白质,2000 年 Taylor 等[111]继续采用 Triton X-114 抽提 CHX SGF1 组分将其分为 Triton X-114 不可溶(insoluble)、水相(aqueous)和去垢剂相(detergent phase)3 个组分,并经二乙基氨基乙基(diethylaminoethyl,DEAE)离子交换层析技术分离富集各组分蛋白质,经双向凝胶电泳(2-DE)分析后,将 CHX SGF1 组分 2-DE 主图谱上检测到的蛋白质点编号由 173 个增加到 588 个。其中丰度最强的 96 个蛋白质点胶内酶解后经串联质谱(MS/MS)分析后鉴定出其中 59 个蛋白质,30 个确定为高尔基体居留蛋白质(如 Giantin、GP25L2、SNARES、Rab、Rap、Rho 家族蛋白等),6 个为内质网循环蛋白质(如 cyclophilin B、BiP 等),7 个为血清蛋白质(如 serum paraxonase 等),6 个为其他细胞器蛋白质(如 Urate oxidase、cathepsin D 等),10 个为胞质蛋白质(如 HSP70、alpha-SNAP 等);另有 22 个蛋白质点也产生有效的质谱图但经检索为未知蛋白质,其中 20 个具有相应的 EST 数据。至此,建立了第一份哺乳动物高尔基体蛋白质组学数据库。

三、蛋白质组学研究发现新的高尔基体蛋白质

Wu 等[115]根据未知蛋白质 GMx33 的肽序列对应的 EST 数据设计出随机引物自大鼠 λ cDNA 文库筛选出两个克隆,并拼接出全长序列。根据可读框预测的蛋白质等电点和相对分子质量与双向凝胶电泳(2-DE)胶上的蛋白质点一致。通过与酵母及各种哺乳动物表达序列标签(EST)数据库比较分析表明 GMx33 有 α、β 两种亚型。免疫印迹发现 GMx33α 在 2-DE 胶上有多种修饰形式,其中有的是磷酸化。Northern 印迹分析表明 GMx33α 存在于大鼠几乎各种器官组织中。GFP-GMx33α、GFP-GMx33β 以及 GFP-酵母融合蛋白质被发现与 TGN46 共定位于高尔基体,免疫荧光和免疫电镜技术进一步证实 GFP-GMx33α 与 TGN38 共定位,从而证明 GFP-GMx33α 分布于反面高尔基体(trans-face of the Golgi complex)。另外,Bell 等[116]结合差速离心和蔗糖密度梯度离心获得高度富集的高尔基体组分并利用 Triton X-114 抽提分离出高尔基体膜蛋白组分后,采用单向 SDS-PAGE 分离。通过延迟抽提基质辅助激光解吸离子化质谱技术(delay extraction matrix-assisted laser desorption ionization mass spectrometry)检测肽质量指纹图谱(peptide mass fingerprinting)、串联质谱技术产生肽序列标签(peptide sequence tag)和 Edman 降解法测蛋白质氨基端序列等方法相结合,共鉴定了 81 种蛋白质,包括 KDEL 受体、p24 家族蛋白、SNARE、Rab、SAMP 等高尔基体居留和转运功能蛋白质,并用免疫胶体金电镜技术验证了 SCAMP、两个 SNARE 及 GBF1、BAP31、α_2P24 等蛋白质在高尔基体内及细胞各分析组分中的分布。其中鉴定了一 34 000Da 新蛋白质 GPP34 并将其分布定位于高尔基体。

四、不同功能状态高尔基体的比较蛋白质组学研究

已知高尔基体在细胞由基础分泌状态转换到最大分泌状态时有显著的形态学和体积

上的改变。例如，乳腺上皮细胞在分娩前后由怀孕末期基础分泌状态到哺乳期最大分泌状态时，高尔基体所占细胞容积由 1%～3%增加到 5%～15%，高尔基体中小池总数增加 2 倍以上[117,118]。这些形态学上的改变与合成酪蛋白(casein)、乳糖(lactose)及脂类等的酶活性增加有关[119]。为了更全面、深刻地理解这些改变的分子机制，Wu 等[120]对乳腺上皮细胞这两种功能状态的高尔基体进行了比较蛋白质组学研究。雌性 SD 大鼠按怀孕期第 18 天(P18)(基础分泌状态)、哺乳期第 10 天(L10)(最大分泌状态)以及是否放线(菌)酮处理分为 4 组，每组 12 只，以消除个体差异，每组大鼠腹股沟乳腺组织剪碎后经胶原酶和透明质酸酶处理释放出乳腺小体(acini)，从富集的乳腺小体中分离出高尔基体，进一步经 Triton X-114 抽提和高碱性碳酸盐(high pH carbonate)处理技术分别获得高尔基体膜蛋白质组分和可溶性蛋白质组分。各个组分蛋白质经双向凝胶电泳分析后表明，高尔基体由基础分泌状态转为最大分泌状态时，全蛋白质中只有部分蛋白质上调。胶内酶解后经串联质谱技术鉴定出 30 种上调蛋白质，其中有 6 种膜内蛋白质(膜联蛋白 Ⅰ，Ⅲ，Ⅳ，Ⅴ，α-SNAP 和突触结合蛋白)、6 种调节蛋白质(Rabs 2，6，11b，11e，14，18)，4 种微管运动蛋白(动力蛋白轻链和中间链，驱动蛋白和肌球蛋白重链)、4 种肌动蛋白调控蛋白(rhoA，rhoGAP，cdc42 和 hsp27)。膜联蛋白和突触结合蛋白调控钙依赖性膜融合(fusion)，而 α-SNAP 在钙依赖性细胞外排(exocytosis)中发挥作用。高尔基体内分泌小泡的移动依赖于微管结构，而其融合、外排活动需要肌动蛋白的重排。这 4 类蛋白质的上调与观察到的高尔基体在最大分泌状态时分泌小泡增多及融合、外排活动增加相一致。

五、总　　结

近 5 年来，高尔基体蛋白质组学研究已经获得大鼠高尔基体及其不同组分的 2-DE 图谱，质谱鉴定出其中百余种蛋白质包括新蛋白 GPP34、GMx33 等。不同功能状态高尔基体的比较蛋白质组学研究揭示了部分蛋白质的表达上调现象。随着分离纯化技术及蛋白质组学技术的发展，对不同生物体及不同生理、病理状态高尔基体的蛋白质组学研究将会进一步促进对高尔基体结构和功能的理解。

第四节　核孔复合体蛋白质组学研究进展

核孔复合体(nuclear pore complex，NPC)是真核细胞核质物质交换的惟一通道。随着对其研究的不断深入，已经提出多种核孔复合体结构模型[121]。了解核孔复合体所有蛋白质组成，对于认识核孔复合体的结构与功能机制有着重要意义。

一、核孔复合体的结构和功能

核孔复合体位于内外核膜彼此融合的区域。随着对其研究的不断深入，其结构模型得以不断完善。Ris[122]利用高压透射电镜技术(HVTEM)和高分辨低压扫描电镜技术(HR-LVSEM)首次揭示了核孔复合体精细的三维结构，发现核孔复合体是一个直径为 120nm 的胞质环(cytoplasmic ring)和核质环(nucleoplasmic ring)所组成的滴漏(hour

glass)样的结构。该模型认为胞质环与核质环构成核孔的外壁,分别与内外核膜相连。从每个环各伸出 8 条丝状物。胞质一边,丝状物短而卷曲;核质一边,丝状物细长,伸入核内 50～70nm,在丝状物末端形成一个直径 60nm 的由 8 个颗粒构成的小环,形成核篮(nuclear basket)样结构,在核孔复合体内部还有一个由 8 个颗粒组成的结构(相当于辐)及一个中央颗粒(相当于栓)。Akey[123]根据各种电镜技术分析核孔复合体二维和三维结构信息,认为核孔复合体对垂直于核膜的中心轴呈八重对称,而对平行于核膜的平面则是两侧不对称的,这与其功能不对称性是一致的,推测核篮和胞质丝状物可能介导运输物与核孔复合体之间的相互作用。这一结构模型特点几乎为各种真核生物所共有,而且许多核孔蛋白(nucleoporins, nups)在不同物种间高度保守[124,125]。

在真核细胞中,核质之间存在着连续而有选择性的双向物质交换,核孔复合体在功能上可被看作是一种特殊的跨膜运输蛋白质复合体,构成核质间双向运输的亲水性通道。其运输方式有两种:被动扩散与主动运输。核孔复合体作为被动扩散的亲水性通道,其功能直径为 9～10nm,有的可达 12.5nm。它允许离子、水溶性分子穿梭于核质之间,进行自由扩散。通过核孔复合体的主动运输包括亲核蛋白质的核输入以及 RNA 及核糖体亚单位的核输出等[121]。作为核质间运输的惟一位点,核孔复合体在细胞周期循环、基因表达调控以及癌症发生等方面都发挥着重要作用[126~128]。

核质间物质交换依赖于被运输物、运输因子(transport factor)及核孔复合体三者间的相互作用。核孔复合体的丝状物上即由一系列含 Phe-Gly 重复序列的核孔蛋白(FG nups)形成结合位点,运输因子在转运信号调控下携带被运输物与这些位点结合,以能量依赖的方式完成物质交换[129~132]。近年来已对核孔复合体的组成进行了一系列研究,发现核孔复合体主要由蛋白质组成[125],但所发现的蛋白质在多大程度上代表了核孔复合体的所有蛋白质组成,以及核孔蛋白质如何协作完成物质转运功能,依然是一个谜[132,133]。对核孔复合体的蛋白质组学研究,将对其结构和功能的理解有着重要意义。

二、酵母核孔复合体蛋白质组学研究

Rout 等于 2000 年[21]将通过不同缓冲液体系和蔗糖密度梯度离心方法相结合[134]分离纯化而富集到的酵母核孔复合体组分蛋白质经羟磷灰石柱和 C4 反相柱高效液相色谱分别分离收集了 80 和 180 个组分。所有组分经 SDS-PAGE 分离、铜染后,切下感兴趣的条带进行胶内胰酶酶解。酶解肽段混合物经 DE-MALDI-TOF MS 或 MALDI ion trap MS 分析、搜索 NR 蛋白质库鉴定出 174 种基因产物,其中 34 个对应于未知 ORF。这些基因产物先根据数据库中的信息粗分为三类:已知 NPC 相关蛋白质(nups 和转运因子)、已知功能明显与核质交换无关的蛋白质以及功能未知蛋白质。研究者重点分析了功能未知的蛋白质。首先通过表达融合蛋白及免疫荧光技术验证了这些蛋白质在核孔复合体上的定位情况;又通过 Nup120p 基因缺失实验将 NPC 与 NE(nuclear envelop)区分开来,排除了 NE 蛋白质;最后根据各种融合蛋白质在亚细胞核组分(NE、nucleoplasmic、cytoplasmic fraction)中的分布情况区别了 Nups 与转运因子。经过这种严格分析,最终确定有 40 种基因产物与 NPC 相关,29 种为 nups,其中 Nup60p/Yar002p、Pom34p/Ylr018p、Cdc31p、Rip1p/Nup42p、Nup192p/Yjl039p、Gel1p 和 Gel2p 为新确定的 nups;11 种为转运因子。免

疫荧光及免疫电镜定位信息表明，绝大多数 nups 对称分布于胞质核质两面；两个 FG nups (Nup1p 和 Nup60p)分布在 NPC 的核质面；而另两个 FG nups (Nup159p 和 Nup42p)和一个非 FG nup(Nup82p)分布在 NPC 的胞质面。这与丝状物上存在的结合位点吻合[68~71]。利用定量免疫印迹技术进行的化学计量学分析表明，NPC 中各种 nups 的丰度可分为 3 类：单面分布的 5 种 nups 中 4 种(Nup1p, Nup60p, Nup42p, Nup159p)为低丰度；大多数 nups 为中等丰度；极少数 nups 为高丰度。

该研究所确定的约 29 种 nups 远低于以前的估计[135]。这可能与由于对称性分布，许多 nups 以多拷贝形式存在于 NPC 有关；另外，从经济角度而言，NPC 可能共享有一些其他细胞组分的蛋白质。按单拷贝低丰度 nups、双拷贝中等丰度 nups 以及四拷贝(或更高)高丰度 nups 计算，分离的 NPC 分子质量约为 44×10^6Da。这一分子质量低于液体动力学方法所估计的 $56\times10^6\sim72\times10^6$Da，但与 STEM 所预测的范围相符[21,136]。根据所鉴定的 NPC 相关蛋白质的种类和定位信息，Rout 等提出了酵母 NPC 的一种结构模型，并进一步对 NPC 的物质交换机制进行了探讨，认为经 NPC 的核质交换可能由结合(docking)、交换(exchange)、捕获(trapping)和释放(vectorial Release)等过程完成[134]。

三、哺乳动物核孔复合体蛋白质组学研究

Cronshaw 等于 2002 年[137]对哺乳动物大鼠核孔复合体 NPC 进行了蛋白质组学研究。大鼠肝细胞核经 DNase 和 RNase 消化以及肝磷脂抽提除去染色质；再经 Triton X-100 和 SDS 处理除去核膜及与核膜相关组分。电镜以及标志性蛋白质的检测证明了所纯化核孔复合体(NPC)的完整性。最后采用两性离子变性剂 Empigen BB 选择性地溶解抽提出包埋于核纤层(lamina)中的 NPC 的 nups，并进一步经 C4 反相柱分离。各个组分经 SDS-PAGE 分离考染，切下条带酶解后进行了 MALDI-QqTOF-MS 或 MALDI-ion trap-MS/MS 分析。通过依次搜索大鼠蛋白质库、人和小鼠蛋白质库以及 EST 数据库，共鉴定了 94 个基因产物，根据文献以及数据库中信息分为 23 个核孔蛋白(nups)、18 个 NPC 相关蛋白质、42 个非 NPC 蛋白质和 11 个功能未明蛋白质。免疫荧光共定位实验证明其中 6 个功能未明蛋白质(MP-44、Sec13L、FLJ12549、p37、p42 和 ALADIN)定位于 NPC。蛋白质氨基酸序列分析表明，MP-44、Sec13L、FLJ12549 分别同已知的酵母核孔蛋白质 Nup53p、Sec13p、Nup85p 有一定的同源性。一部分已知 nups 具有的 FG 重复序列是核质物质交换中转运因子的结合位点[138]，但在 p37、p42 和 ALADIN 中发现它们都具有 WD 重复序列。根据各个 nup 的相对丰度估计其在每个 NPC 中的拷贝数，并进一步预测出单个 NPC 的分子质量约 60×10^6Da。

对哺乳动物大鼠核孔复合体蛋白质组学研究，表明哺乳动物的核孔复合体也是由约 30 种核孔蛋白组成，并发现了一类新型的具有 WD 重复序列的核孔蛋白。

四、总　　结

结构学比较表明，脊椎动物核孔复合体较酵母的大，分别为 145nm×80nm 和 96nm×35nm[136,139]。分子质量测定也表明脊椎动物核孔复合体的 125×10^6Da 分子质量远大

于酵母的 $55\times10^6\sim72\times10^6$Da[21,136,140]。这也与蛋白质组学研究所估计的哺乳动物 60×10^6Da 的 NPC 较酵母的 44×10^6Da 大致吻合。但对酵母及哺乳动物核孔复合体的蛋白质组学研究表明，它们都只由约 30 种 nups 组成，这主要是与多拷贝高分子质量的核孔蛋白形成核孔复合体的对称性结构有关。而且约 2/3 的 nups 在二者间是保守的，这表明哺乳动物与酵母的核孔复合体虽然存在物种间差异表现出的大小和组成上的差异，在结构与功能上仍有进化上的保守性。

对酵母和哺乳动物核孔复合体的蛋白质组学研究分别发现了 7 种和 6 种新核孔蛋白质，发现了一类新型的具有 WD 重复序列的核孔蛋白质，并且根据研究结果对核孔复合体的分子质量、分子结构模型和核质交换功能的机制进行了深入探讨，其重要意义已不言而喻。Blobel 等[141]则认为 Rout 等人的工作不仅对核孔复合体的分子结构提供了大量信息，研究方法可被对其他大分子复合体的研究借鉴；而且引发了对蛋白质和蛋白质相互作用以及分子运输机制的深入思考。

第五节　核仁蛋白质组学研究进展

核仁(nucleolus)是真核细胞核内一动态组织结构。除核糖体亚基的生物生成功能外，已发现还具有其他多种生物功能。核仁的蛋白质组学研究将促进深入理解核仁的细胞周期性的动态变化及具有多种生物功能的机制。

一、核仁的结构与功能

核仁是一亚细胞核组织结构，其存在部位与一些特殊染色体上存在的含有 rRNA 基因的核仁组织区(nucleolar organizer region，NOR)密切相关。核仁是一动态结构，在真核细胞间期核中结构最明显，在每个细胞周期中经历解聚与重建的过程[142]。电镜下观察核仁的超微结构，可辨认出 3 个特征性的区域：纤维中心(fibrillar center，FC)、致密纤维组分(dense fibrillar component，DFC)和颗粒组分(granular component，GC)。一般认为，rDNA转录局限在 FC 周边部分，促使 rRNA 转录物集聚于 DFC 进行修饰处理，而在 GC 中组装成核糖体亚单位[143~146]。

核仁最主要的功能是核糖体亚基的生物生成，包含 rRNA 的合成、加工和核糖体亚单位的组装等过程。最近研究表明核仁还具有其他功能[147,148]，如其他核蛋白体，包括信号识别颗粒(signal recognition particle，SRP)[149]、剪接体中的小核蛋白体[150]和端粒酶[151]等的生成或成熟场所；核仁还参与一些 mRNAs 和 tRNAs 的转录后处理以及核输出[152,153]；并且通过扣押(sequestration)机制调控一些特殊调节性蛋白质的活性等[154]。

核仁中主要含蛋白质，为核仁干重的 80%；RNA 为核仁干重 10%左右；DNA 含量较 RNA 少；脂类含量极少。细胞化学研究发现核仁中存在碱性磷酸梅、ATP 酶等多种酶系[155]。对核仁中各种组分特别是蛋白质组成的了解将促进对核仁结构及多种功能的认识。

二、核仁的蛋白质组学研究进展

Andersen 等人[156]联合应用超声波和蔗糖密度梯度离心技术分离纯化出 HeLa 细胞的核仁。通过光镜、电镜检查以及标志性蛋白质检测验证了核仁的纯度和完整性后，核仁组分蛋白质分别经 Tris-甘氨酸、Tris-乙酸盐 1D SDS-PAGE 和 2D SDS-PAGE 获得最大程度分离。凝胶上的条带和蛋白质点切下酶解后应用 MALDI-TOF-MS 或 nanoES-MS/MS 分析。通过逐级搜索人 NR 蛋白质数据库、EST 数据库以及人类基因组测序草图，共鉴定 271 个基因产物，其中 30.5% 为新基因产物、22% 为核苷酸或核酸结合蛋白质、14% 为核糖体蛋白质、9% 为 RNA 修饰酶或相关蛋白质、5% 为含具有 RNA 依赖的 ATP 酶和螺旋酶活性的 DEAD box 的蛋白质、5% 为分子伴侣蛋白质，另外还有 3.5% 的翻译因子。选择构建、表达 18 个新基因产物的融合蛋白，经免疫荧光学检测验证，其中 15 个在核仁中有分布，这表明所鉴定的基因产物绝大部分在体内情况下与核仁有关。

为了了解核仁结构及蛋白质组成的动态性，Andersen 等人[156]采用转录抑制性药物 Actinomycin D 处理 HeLa 细胞。已有报道 Actinomycin D 可影响细胞核的结构、核蛋白质的分布模式以及核仁的结构完整性[157]。对处理组和对照组的比较蛋白质组学研究鉴定出 11 种蛋白质的丰度在处理组的核仁中有增加，分别为含 DEAD box 蛋白质（DDX9、p72、p68）、hnRNP 蛋白质（hnRNPs K，G 和 A2/B1）、RNA 结合相关蛋白质（PSF，PSP2/CoAA，PSP1 和 p54/nrb）和 p80 coilin。构建、表达了 p72、p68、PSP2/CoAA、PSP1 和 p54/nrb 的 YFP 融合蛋白质，令人惊异地发现所有这些融合蛋白质在对照组 HeLa 细胞中主要存在于核基质，但经 Actinomycin D 处理后，都转移到核仁中。表明在代谢紊乱情况下，核仁的蛋白质组成有动态性变化。

Fox 等人[158]根据新鉴定出的核仁蛋白质 PSP1 在核仁中的动态分布信息，在各种人细胞类型中发现了细胞核中存在的新的结构区域 paraspeckle，并证明 paraspeckle 至少有 PSP1、PSP2 和 p54/nrb3 种蛋白质组成，而且在转录抑制后，这 3 种蛋白质都转移到核仁周边的帽状结构上。

三、总　　结

Andersen 等人对人 HeLa 细胞核仁的蛋白质组学研究揭示了核仁具有出乎意料的复杂的蛋白质组成，而且随着代谢状态的不同有动态变化。对这项工作新发现的细胞核内的结构区域 paraspeckle 的结构与功能的研究将对分子细胞生物学的发展具有重要意义。Pederson[159]、Schuldt[160]、Dundr[161]、Ospina[162]等均积极评价了此项工作。

第六节　线粒体蛋白质组学研究

线粒体是真核细胞内一重要细胞器，除作为能量生产场所外，已发现其参与多种细胞的生理病理活动。Science 于 1999 年 3 月发表了“Mitochondrial make a comeback”专题系列论文，重申线粒体再度成为当前生命科学和分子医学中的热点和前沿[163]。近年来，蛋

白质组学技术已被广泛应用于多种生物线粒体的研究，并取得许多有意义的成果。我们结合差速离心与双向凝胶电泳技术初步获得大鼠肝组织线粒体蛋白质的双向凝胶电泳图谱，通过与前人工作的借鉴比较，为深入研究亚细胞蛋白质组学做了一定的准备。

一、线粒体蛋白质组学研究进展

（一）线粒体的结构与功能

线粒体的外形多种多样，如椭圆形、哑铃形、环形、圆柱形，与细胞类型和所处生理条件有关，最常见的是呈线状或颗粒状。线粒体的大小随细胞不同类型而异，一般直径为0.5～1μm，长2～3μm。线粒体的数目在不同类型细胞中差异很大，哺乳动物成熟的红细胞中没有线粒体，一般动物细胞中平均有数百个，如大鼠肝细胞内的线粒体数为500～1 400个；但在卵母细胞中可高达300 000个。一般地，植物细胞中线粒体含量比动物细胞要少，可能与富含叶绿体有关。线粒体的数目与细胞的生理功能及需求密切相关，一般而言，需要能量较多的细胞，线粒体数目也较多。在电镜下观察，线粒体有着相同的结构，它是一个由内外两层单位膜构成的封闭的囊状结构，主要由外膜、内膜、膜间隙和基质（或称内室）组成[164]。

线粒体的化学组成[164]主要是蛋白质，其次是脂类，另外还有许多辅酶、维生素、金属离子和线粒体DNA、RNA、核糖体等。蛋白质含量占线粒体干重的65%～70%；蛋白质在线粒体各组成部分分布有很大差异。以大鼠肝线粒体为例：67%蛋白质在基质中，21%在内膜，6%在外膜，6%在膜间隙。线粒体蛋白质可分为两类：可溶性和不溶性。可溶的蛋白质大多数是基质中的酶和一定数量的外周膜蛋白；不溶性蛋白一般是构成膜的镶嵌蛋白、结构蛋白和部分酶蛋白。线粒体中已被确认的酶有上百种，分布在各个结构组分中。如外膜中含有合成线粒体脂类的酶类，内膜中含有执行呼吸链氧化反应的酶系和ATP合成酶系；基质中有高浓度的多种酶的混合物，如参与三羧酸循环反应、丙酮酸与脂肪酸氧化的酶系和蛋白质与核酸合成酶类等。Jung等[36]利用PSORT软件通过分选信号（sorting signal）的检测将SWISS-PROT、TrEMBL、TrEMBL-NEW蛋白质数据库中的共24 846个人类蛋白质进行了亚细胞定位预测，发现其中约有3 362个线粒体蛋白质。说明尚有许多未知线粒体蛋白质有待确认。

线粒体不仅维持细胞正常生理功能，在细胞凋亡和死亡中也发挥重要作用。线粒体是细胞能量生成场所，通过氧化磷酸化为有机体提供90%的ATP能源；参与了尿素循环。它调节细胞内Ca^{2+}平衡，从而参与细胞信号传导[165～167]。线粒体参与细胞凋亡功能的发现加深了对细胞凋亡机制的理解[168～170]。它还产生还原性物质有利于细胞的解毒功能[171]。可见，线粒体在生物体的生长、发育、代谢、衰老、疾病、死亡以及生物进化等方面都有非常重要的意义。

（二）线粒体的分离纯化

线粒体分离的基本步骤是，先制备组织或细胞匀浆，然后进行差速离心；如果有必要，还可进行密度梯度离心而进一步纯化。组织细胞的特异性及所研究线粒体的实验用途决

定了方法的细节。这些细节包括:匀浆器的选择、缓冲液的成分、许可混杂其他细胞器的程度等。通常用 Ficoll、Percoll、碘化介质(OptiPrep,Nycodenz,metrizamide)和蔗糖来形成密度梯度进一步纯化线粒体。

Ficoll 是最早发展起来的专门用于细胞分级分离的介质[172,173]。它最常用于非连续梯度,普遍用于从脑部分离线粒体和神经突触体。对于 Percoll(聚乙烯吡咯烷酮包被的硅颗粒),可使用预先形成的不连续或连续梯度,或在离心过程中使其自动形成梯度,线粒体可以在这个梯度中形成等密度带,其他细胞器因密度较低而位于离心管上部。它广泛用于各种细胞器的分离。其整个梯度都是等渗的,并且能很快地自动形成梯度。然而作为一种胶体悬液,在使用中存在着较难清除的缺点。使用 Nycodenz 和 metrizamide 时,要预先形成连续或不连续梯度。Nycodenz 能自动形成梯度并已用于过氧化物酶体的纯化,但该过程较为缓慢。OptiPrep(Iodixnol)是 Nycodenz 的二聚体,具有短时间(约 3h)内自动形成梯度的优势。像 Percoll 一样,其整个梯度都是等渗的。在 OptiPrep 中分离线粒体[174]是比较令人感兴趣的,因为在 1 000 000g 离心时形成一个较窄的梯度,过氧化物酶体会在底部形成沉淀,而溶酶体和其他细胞器浮在梯度顶端的浸液部分,线粒体则分布于梯度的较浅部分。因为不形成厚的细胞器带,不会带来被缠绕的问题。移去梯度顶层后,收集留下的包括线粒体的梯度部分并离心沉淀。若有必要,还可通过稀释和漂洗线粒体或透析的方式除去 OptiPrep。非连续蔗糖梯度是最常用的介质,尽管线粒体会被离心时产生的高渗透压破坏,一些代谢功能也会改变或丧失。因为蔗糖比较容易得到,且较为便宜,得到的线粒体可在许多实验中使用。

(三)线粒体的蛋白质组学研究

对不同生物体线粒体的蛋白质组学研究已有许多报道。Rabilloud 等于 1998 年通过双向凝胶电泳技术分离人胎盘线粒体蛋白质,并主要采用质量肽谱法,结合 N 端测序和免疫印迹法鉴定出 50 余种人胎盘线粒体蛋白质[12]。

大鼠肝脏中富含线粒体。用不同方法从不同角度对大鼠肝组织线粒体的蛋白质组学研究也多有报道。2000 年 Lopez[19]等使用高通量自动化设备结合微旋转亲和柱亲和层析技术对差速离心法获得的雄性 Fischer 344 大鼠的肝脏线粒体进行了蛋白质组学研究,数天之内鉴定了百余种线粒体蛋白质,并通过钙离子螯合亲和、ConA 树脂亲和、苯基琼脂糖树脂亲和技术分别富集了大鼠线粒体的钙结合蛋白质、糖蛋白和膜蛋白,获得了相应的 2-DE 图谱。2001 年 Murayama[31]等人利用 Nycodenz 密度梯度离心一步法分离出成年 Wistar 大鼠肝脏组织的各种细胞器,经标志酶检测鉴定后,对各种细胞器(包括线粒体)组分进行了初步的 2-DE 分析。2002 年 Fountoulakis[83]等人差速离心法获得成年雄性 Wistar 大鼠肝脏线粒体后,采用 pH3～10 非线性,pH4～7 线性,pH5.5～6.7 线性 3 种 IPG 胶条对其进行第一相 IEF 电泳分离,继以 10%SDS-PAGE 进行第二相分离。对 6 块 2D 凝胶上的 1 800 个蛋白质点进行 MALDI-MS 鉴定,检测到 192 种不同的基因产物。其中约 42%的蛋白质点已被亚细胞定于线粒体。有 8 种基因产物为首次检测且均由一个蛋白质点代表。

2001 年 Kruft 等[175]为全面考察植物线粒体的功能,对拟南芥线粒体进行了蛋白质组学研究。他们通过 IEF/Tricine SDS-PAGE 双向凝胶电泳技术分离拟南芥线粒体蛋白

质，银染重复显示出650个不同蛋白质点，并用免疫印迹、直接蛋白质测序和质谱法随机鉴定了52个蛋白质点，进行了功能分类。与此同时，Millar[176]等也发表了对拟南芥线粒体蛋白质组学研究的论文，双向凝胶电泳分离染色显示了约100种高丰度蛋白质和250种低丰度蛋白质，并将线粒体蛋白质样品按膜关联程度进一步分级，分别得到全部蛋白质、可溶性蛋白质、膜蛋白和内在膜蛋白质的2D图谱。实验共定位了163种蛋白质，其中91种通过搜索拟南芥基因组翻译数据库得到鉴定。

（四）线粒体的三维分离及线粒体亚组分的蛋白质组学研究

以上获得的线粒体的2D图谱大多是溶解整个线粒体乃至整个细胞所得，这使得2D图谱含有的点过多，反而不利于分析。而且这类图谱也难以提供单个多蛋白质复合体的蛋白质组成信息和功能信息。

Hanson等[177]将蔗糖密度梯度离心与双向凝胶电泳技术结合，形成线粒体蛋白质的三维分离，其中蔗糖密度梯度离心可按大小分离线粒体中多蛋白质复合体。该法在牛心脏、MRC纤维原细胞和人脑组织的线粒体的应用结果表明有助于解决等电聚焦过程中蛋白质的分辨问题以及疏水蛋白质的溶解问题等，并可提供线粒体内多蛋白质复合体的结构功能信息。

2002年Werhahn[178]等也发展了一种3D电泳，首先用Blue-native PAGE分离线粒体多蛋白质复合体，再电洗脱至完全解离考马斯亮蓝染料，沉淀后继以IEF/SDS-PAGE分离多蛋白质复合体，形成Blue-native/IEF/SDS-PAGE三维电泳。利用该法分离了拟南芥线粒体的ATP合成酶复合物、细胞色素C还原酶复合物和外膜移位酶（TOM）复合物，并将获得的各复合物的蛋白质组图谱与线粒体整体蛋白质组图谱进行了比较。三维分离技术对线粒体组分再次细分，进一步富集了低丰度蛋白质，提高了蛋白质组图谱的分辨率；并使分离鉴定线粒体中多蛋白质复合体的蛋白质组成成为可能，促进了对所有线粒体蛋白质组成及其功能的研究。

除三维分离技术外，不少学者也通过其他各种方法对线粒体的多蛋白质复合体亚组分进行了蛋白质组学研究。线粒体外膜移位酶（TOM）复合物是一种介导核编码的线粒体蛋白质穿越线粒体外膜的移位酶，对其亚基组成的鉴定将有助于了解线粒体内外蛋白质的转运机制。Jansch等[179]和Werhahn等[180]以Blue-native PAGE/Tricine-SDS-PAGE分别分离了马铃薯和拟南芥TOM复合物，前者分离得到70×10^3、36×10^3、23×10^3、9×10^3、8×10^3、7×10^3、6×10^3Da 7种不同亚基，后者得到34×10^3、23×10^3、21×10^3、8×10^3、7×10^3、6×10^3Da 6种不同亚基。通过免疫印迹和直接蛋白质测序鉴定了其中某些蛋白质，并与酵母TOM复合物亚基蛋白质进行了同源性比较，提出了这些蛋白质可能的功能。

核糖体是合成蛋白质的细胞器，由大小两个亚基组成。真核细胞线粒体核糖体的沉降系数常较同一生物体的细胞质核糖体（80S）低而近似于原核细胞的核糖体（70S）。哺乳动物线粒体核糖体更低至55S[181,182]，由39S大亚基和28S小亚基组成，与细菌核糖体有相似的分子质量（2.8×10^6Da）[183]。对线粒体核糖体结构功能的研究对理解线粒体的生物合成功能以及线粒体基因组的系统发生起源具有重要意义。蛋白质组学技术方法为线粒体核糖体的结构功能研究提供了强有力的工具。1995年Cahill[30]等人通过双向凝

胶电泳技术分离到86种与雄性SD大鼠肝线粒体核糖体相关蛋白质、70种与细胞质核糖体相关蛋白质，并通过2DE图谱比较发现细胞质核糖体蛋白质较线粒体核糖体蛋白质偏碱性。近两年来，Koc等[184,185]利用双向凝胶电泳技术分离牛线粒体核糖体28S小亚基蛋白质，选取某些蛋白质点胶内胰酶消化后，采用液相色谱电喷雾串联质谱(LC-ESI-MS/MS)分离肽段测定肽序列，经搜索EST数据库共鉴定了10种新的不同类型的核糖体小亚基蛋白质，使被鉴定的核糖体小亚基蛋白质由5种增至15种。经同源性比较，其中MRP-S5,9,10,14与原核生物的核糖体蛋白质有明显的序列相似性。并发现MRP-S25为低等和高等真核生物线粒体共有而在原核生物中却未发现的一种核糖体蛋白质。Suzuki等[186]以无自由基的高还原性(radical-free high-reducing,RFHR)双向凝胶电泳(2D-PAGE)分离牛线粒体核糖体28S小亚基蛋白质，经LC-ESI-MS/MS进行分析，通过搜索ESI数据库鉴定了含11种新蛋白质在内的21种小亚基蛋白质，其中包括S12蛋白(一种关键调节蛋白质)的一种线粒体同源物。此外，还发现凋亡相关蛋白质DAP3是核糖体小亚基的一种成分，表明线粒体核糖体在程序性死亡中具有新的作用。Suzuki等[187]还利用同样的方法分离鉴定了31种牛线粒体核糖体大亚基蛋白质，其中15种为新蛋白质。并表明线粒体核糖体大亚基中增大的蛋白质对rRNA缺失具有结构补偿作用。

（五）线粒体的蛋白质组学差异图谱分析

线粒体作为真核细胞内一重要细胞器，在许多病理状态下或药物、毒物作用下，常有结构或功能上的变化或损伤。采用蛋白质组学方法对线粒体在不同状态下的蛋白质组图谱进行差异显示分析研究，可促进对发病机制及药理、毒理作用的理解。

线粒体功能紊乱与阿尔茨海默病(Alzheimer's diseases,AD)及唐氏综合征(Down syndrome,DS)相关。Kim等[188]用蛋白质组学方法对AD病和DS病成年病人脑区线粒体的蛋白质的差异表达情况进行了研究，发现CoQ:细胞色素c氧化还原酶(复合物Ⅲ)核心蛋白1在AD病人的脑颞叶中明显下降，而ATP合成酶(复合物Ⅴ)β链在DS病人的脑额叶中明显下降。这些线粒体呼吸酶的减少可能导致在AD病中观察到的能量代谢损伤，也可能导致DS和AD病人脑中活性氧(ROS)的生成和神经细胞凋亡。Kim等[189]又利用蛋白质组学技术定量考察了DS病胎儿大脑皮层线粒体电子传递链酶蛋白质水平的变化，发现NADH:CoQ氧化还原酶(复合物Ⅰ)的30×10^3Da亚基在蛋白质水平显著下降，推测可能是这种下降导致细胞的能量和自由基代谢受损，继而影响DS胎儿的脑发育。

醋氨酚(acetaminophen,APAP)为临床常用的解热镇痛药，但APAP过量使用可产生不同程度的肝损伤。为探明其毒性机制，Tonge等[190]用荧光双向差异凝胶电泳(2D-DIGE)技术比较了毒性剂量、亚毒性剂量APAP处理后的小鼠与正常小鼠的肝线粒体蛋白质组图谱，发现73种蛋白质的丰度有明显变化。质谱鉴定了其中43种蛋白质，发现75×10^3Da GRP、Hsp10、Hsp60和过氧化氢酶等的变化有剂量和时间依赖性。Ruepp等[191]进一步研究发现大多数线粒体蛋白质变化出现在注射APAP后15min，这要先于大多数的基因表达调控过程，表明APAP的毒性可能并非基因表达调控的结果，其中Hsp10和Hsp60均很快减半，推测很可能是泄露到胞质中。

红藻氨酸(kainic acid,KA)是一种强效的神经毒素和刺激性氨基酸,可导致脑蛋白质的混乱。Krapfenbauer 等[192]用蛋白质组学方法系统地研究了 KA 对雄性 SPD 大鼠脑蛋白质表达水平的影响,发现 KA 处理后线粒体中最明显的变化是微管蛋白 α-1 链部分片段蛋白质点的出现,而在未处理脑线粒体中该蛋白质仅以低水平存在。另外,KA 处理后,NF-L 蛋白(neurofilament triplet L-protein)和 α-internexin 等蛋白质明显降低,葡萄糖调节蛋白 GRP78 也有所下调。二氢硫辛酰胺脱氢酶、ATP 合成酶 β 链和异柠檬酸脱氢酶的表达下降而丙酮酸激酶上调表明可能是线粒体受损所致。

(六)总　　结

总之,随着对线粒体各种功能以及与人类疾病关系的深入研究,线粒体再度成为生命科学和分子医学中的研究热点。近几年来,蛋白质组学在线粒体研究中的应用已取得重大成果,如构建了人、大鼠、拟南芥的线粒体蛋白质组图谱,建立了线粒体蛋白质组数据库 MITOP(http://www.mips.biochem.mpg.de/proj/medgen/mitop/)和拟南芥线粒体蛋白质组数据库(http://www.gartenbau.uni-hannover.de/genetik/AMPP)[73],并对 TOM 复合物、线粒体核糖体等线粒体亚组分进行了蛋白质组学分析。对线粒体在不同病理、药物作用状态下的蛋白质组图谱进行差异显示分析研究,促进了对发病机制及药理、毒理作用的理解。不断发展的蛋白质组学技术在线粒体研究上的应用必将促进对线粒体的结构功能的深入理解。

大鼠肝脏已成为一普遍的生物研究组织材料来源。大鼠肝脏中富含线粒体。对大鼠线粒体的分离技术已经成熟[19,83,193],不同方法从不同角度对大鼠肝组织线粒体的蛋白质组学研究也多有报道[19,31,83]。

二、材　　料

(一)动物来源

180g 雄性 SD 大鼠 1 只,购自复旦医学院实验动物部。

(二)试剂和溶剂

①所有缓冲液均用经过 Millipore (Bedford,MA,USA)系统处理过的水配制。

②甘露醇(mannitol),购自 Amresco(Solon,Ohio,USA)公司。

③蔗糖(sucrose),购自 Bio Rad(Richmond,CA,USA)公司。

④乙二胺四乙酸二钾盐(EDTA),购自 Amresco(Solon,Ohio,USA)公司。

⑤乙二醇双(2-氨基乙基)醚-*N*,*N*,*N*′,*N*′-四乙酸(EGTA),购自 Amresco(Solon,Ohio,USA)公司。

⑥苯甲基磺酰氟(PMSF),购自 Amresco(Solon,Ohio,USA)公司。

⑦*N*-2-羟乙基哌嗪-*N*′-3-丙磺酸(HEPES),购自 Amresco(Solon,Ohio,USA)公司。

⑧牛血清白蛋白(bovine serum albumin),购自 Sigma (St. Louis,MO,USA)公司。

⑨尿素(urea),购自 Bio Rad(Richmond,CA,USA)公司。

⑩imethylammonio-1-propanesulfonate(CHAPS),购自 Bio Rad 公司。
⑪硫苏糖醇 (DTT),购自 Bio Rad(Richmond,CA,USA)公司。
⑫代乙酰胺 (IAA),购自 Bio Rad(Richmond,CA,USA)公司。
⑬N,N,N',N'-四甲基乙二胺,购自 Bio Rad(Richmond,CA,USA)公司。
⑭甘油(glycerol),购自 Bio Rad(Richmond,CA,USA)公司。
⑮十二烷基硫酸钠(SDS),购自 Bio Rad(Richmond,CA,USA)公司。
⑯三(羟甲基)氨基甲烷(Tris),购自 Bio Rad(Richmond,CA,USA)公司。
⑰硫代硫酸钠(sodium thiosulfate),购自 Amresco(Solon,Ohio,USA)公司。
⑱乙酸钠(sodium acetate),购自 Amresco(Solon,Ohio,USA)公司。
⑲碳酸钠(sodium carbonate),购自 Amresco(Solon,Ohio,USA)公司。
⑳甲醛(formaldehyde),购自 Sigma (St. Louis,MO,USA)公司。
㉑Immobiline Drystrip pH3～10NL 13cm pH4～713cm,购自 Amersham Pharmacia Biotech (Uppsala,Sweden)公司。
㉒两性电解质 Pharmalyte pH3～10NL 和 pH4～7L,购自 Amersham Pharmacia Biotech (Uppsala,Sweden)公司。

(三)仪　　器

肝组织匀浆在电机制动的 Teflon 匀浆器上进行,差速离心使用 Beckman(Fullerton, CA,USA)CS-15R 离心机。超声使用宁波新芝科器研究所 JY92 超声波细胞粉碎机。双向凝胶电泳第一相等电聚焦在 IPGphor™ Isoelectric Focusing System(Amersham Pharmacia Biotech, Uppsala, Sweden)上进行,第二相 SDS-PAGE 在 Hoeffer Electrophoresis Unit SE600(Amersham Pharmacia Biotech, Uppsala, Sweden)上进行。

三、方　　法

1. 差速离心法制备大鼠肝组织线粒体

取 180g 雄性 SD 大鼠 1 只(购自复旦医学院实验动物部),断头法处死后,迅速解剖取出肝脏,将约 1 g 肝组织放入 10ml 冰浴过的等渗溶液 Buffer A(250mmol/L 甘露醇, 75mmol/L 蔗糖, 0.1mmol/L EDTA, 0.5mmol/L EGTA, 1mmol/L PMSF, 10mmol/L HEPES, pH 7.4)中,用电机制动的 Teflon 匀浆器上下 5～8 次进行匀浆。匀浆液 1 000g 10min 离心两次,去除细胞膜及细胞核成分,移出上清进一步 10 000g 离心 15min。沉淀便是肝组织线粒体成分,用 10ml Buffer B(Buffer A 加 0.5%牛血清白蛋白)洗两次,加 1ml 裂解液 (8mol/L 尿素, 4% CHAPS, 65mmol/L DTT, 40mmol/L Tris)裂解后, 100 W4 ×15 S 超声。4℃ 25 000g 离心 1h(Beckman 仪器, F2402 转头, 15 000 r/min 1h),取上清。采用 Bradford 法定量蛋白质浓度,浓度为 7.2μg/μl,总计 7～8 mg 蛋白质。根据上样量分装后 -80℃保存。而上清部分再经 100 000g 离心 1h,沉淀富集微体组分蛋白质,再用 1ml 裂解液裂解,处理过程同线粒体成分,测得浓度为 13.4μg/μl,总计为 13～14 mg 蛋白质。

2. 大鼠肝组织线粒体蛋白质样品的 2-DE 分析

200～400μg 蛋白质样品和重泡涨液(8mol/L 尿素,2% CHAPS,18mmol/L DTT,0.5%pH3～10 两性电解质,痕量溴酚蓝)充分混合,总体积为 250 μl,采用 pH3～10 NL 或 pH4～7 胶条进行第一向等电聚焦。在 IPGphor™ Isoelectric Focusing System(Amersham Pharmacia Biotech)上,重泡涨 30V 12h,500V 1h,1 000V 1h,8 000V 8～10h。整个过程在 20℃进行,电流上限为 70μA。经溶于平衡缓冲液(50mmol/L Tris-HCL pH6.8,6mol/L 尿素,30% 甘油,2% SDS,痕量溴酚蓝)中的 0.8% DTT,2% IAA 分别平衡 15min 后,第一向胶条转移至预先制好的浓度为 12.5%或 10%的 SDS 聚丙烯酰胺凝胶上进行第二向 SDS-PAGE,10 mA/胶走 20min,然后 20mA/胶走至溴酚蓝离胶下沿 0.5cm,全程约 5h 30min。

3. 银染方案

卸胶后经 40% 乙醇、10% 乙酸固定后,进行银染,银染方案为 Yan 等人[194]改进的适用于直接取点做质谱鉴定的染色方案,将 SDS-PAGE 胶 Milli-Q 水漂洗 5min,在含 40%甲醇和 10%乙酸的固定液中固定 15min,共 2 次,在含 5%硫代硫酸钠、68g/L 乙酸钠、30%甲醇中增感作用 30min,水洗 5min 3 次。0.25%硝酸银作用 20min,水洗 1min 2 次。用 25g/L 碳酸钠、0.4ml/L37%甲醛显影 5min 左右至背景出现,14.6g/L EDTA 停显 10min,水洗 5min,共 3 次。该方案灵敏度降低,但比考马斯亮蓝染色仍灵敏得多。

四、结　　果

1. 大鼠肝组织线粒体蛋白质的 2-DE 图谱

由 2-DE 图谱可见,在 pH3～10 和 pH4～7 范围的两块胶上都分离到很多的蛋白质点,经 Bio-Rad PDQuest 软件分析,较明显的蛋白质点有 2 000～2 500 个;线粒体蛋白质在碱性端有富集现象。与 Fountoulakis 等人[83]获得的大鼠线粒体的 2-DE 图谱相比,有着极相似的 pattern,但较考马斯亮蓝染色及银染灵敏得多,染出更多的蛋白质点。

2. 大鼠肝组织线粒体蛋白质与微体蛋白质的 2-DE 图谱比较

通过差速离心分离得到的线粒体与微体组分蛋白质的 2-DE 图谱有着相异的 pattern,达到了分级分离以富集、检测更多蛋白质的目的。但不同组分的纯度很难保证,拟用 Nycodenz 将差速离心获得组分进一步用密度梯度离心法纯化后进行蛋白质组学分析。

第七节　展　　望

刚刚发展起来的亚细胞蛋白质组学研究尽管已经取得了令人瞩目的成果,如不同亚细胞结构的蛋白质组成、新蛋白质的发现甚至新亚细胞结构区域的发现等,从而深入促进了对亚细胞结构与功能的理解;仍然具有广阔的发展前景。

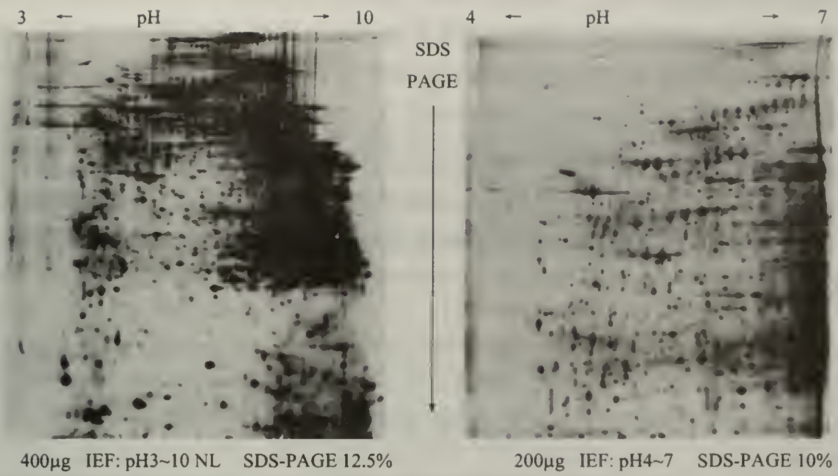


图 19.1 大鼠肝组织线粒体蛋白质双向疑胶电泳银染图谱

左：上样量 400μg，第一向等电聚焦（IEF）pH3～10，第二向 SDS 聚丙烯酰胺凝胶电泳胶浓度 12.5%；

右：上样量 200μg，第一向等电聚焦（IEF）pH4～7，第二向 SDS 聚丙烯酰胺凝胶电泳胶浓度 10%

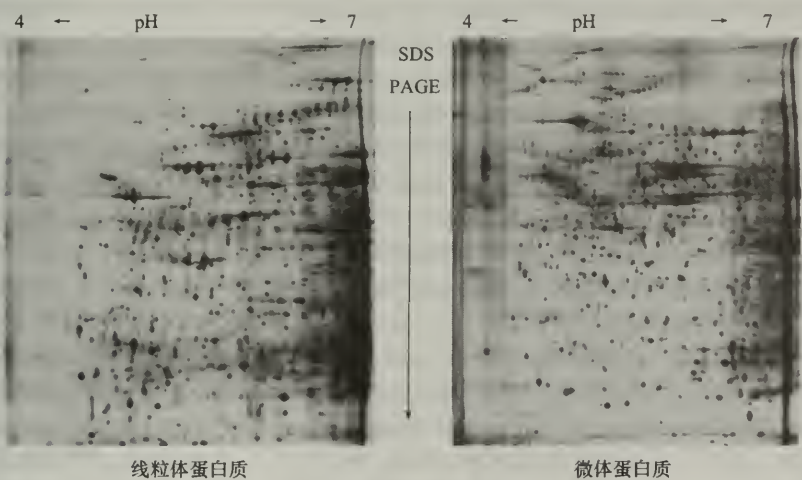


图 19.2 大鼠肝组织线粒体和微体蛋白质双向凝胶电泳银染图谱比较

上样量 200μg，第一向等电聚焦（IEF）pH4～7，第二向 SDS 聚丙烯酰胺凝胶电泳胶浓度 10%。

左：线粒体蛋白质；右：微体蛋白质

作为亚细胞蛋白质组学研究中两大核心技术，亚细胞组分分离技术和蛋白质组学技术，其发展日新月异。除了传统的差速离心与密度梯度离心技术外，其他如自由流电泳技术以及一些分子细胞生物学中的实验技术业已应用于亚细胞组分的分离。分离到的亚细胞组分甚至亚亚细胞组分的种类与纯度都在日益增加。高效液相色谱（HPLC）或多维色谱技术分离蛋白质或肽段混合物与串联质谱技术的联用极大促进了蛋白质组学尤其是亚细胞蛋白质组学研究的发展，可以使多蛋白质复合体性质的亚细胞组分中不同特性的蛋白质得到很好鉴定。不同物种基因组数据库以及蛋白质数据库的不断完善将会鉴定出更

多的亚细胞组分中的蛋白质，发现更多的新蛋白质。

蛋白质组学中定量技术的发展将会促进蛋白质在不同亚细胞组分中差异分布的研究。亚细胞蛋白质组学研究与翻译后修饰蛋白质组学研究的结合将会促进对亚细胞组分中蛋白质翻译后修饰情况的了解，如亚细胞组分中糖基化、磷酸化蛋白质的研究。不同物种间亚细胞蛋白质组学的比较分析，将有助于对物种进化的研究；而生理与病理情况下的亚细胞蛋白质组学差异比较分析，将对疾病的病理机制乃至疾病的诊断治疗等都具有重要意义。

总之，不断发展的亚细胞蛋白质组学的研究与细胞生物学、分子生物学相结合，将会深入促进生命科学的发展。

（江晓胜）

参考文献

1 Wilkins MR, Sanchez JC, Gooley AA *et al*. Progress with proteome projects: why all proteins expressed by a genome should be identified and how to do it. Biotechnol Genet Eng Rev, 1996, 13: 19～50

2 Klose J. Protein mapping by combined isoelectric focusing and electrophoresis of mouse tissues. A novel approach to testing for induced point mutations in mammals. Humangenetik, 1975, 26: 231～243

3 O'Farrell PH. High resolution two-dimensional electrophoresis of proteins. J Biol Chem, 1975, 250: 4007～4021

4 Scheele GA. Two-dimensional gel analysis of soluble proteins. Charaterization of guinea pig exocrine pancreatic proteins. J Biol Chem, 1975, 250: 5375～5385

5 Henzel WJ, Billeci TM, Stults JT *et al*. Identifying proteins from two-dimensional gels by molecular mass searching of peptide fragments in protein sequence databases. Proc Natl Acad Sci USA, 1993, 90: 5011～5015

6 James P, Quadroni M, Carafoli E *et al*. Protein identification by mass profile fingerprinting. Biochem Biophys Res commun, 1993, 195: 58～64

7 Mann M, Hojrup P, Roepstorff P. Use of mass spectrometric molecular weight information to identify proteins in sequence databases. Biol Mass Spectrom, 1993, 22: 338～345

8 Yates Ⅲ JR, Speicher S, Griffin PR *et al*. Peptide mass maps: a highly informative approach to protein identification. Anal Biochem, 1979, 214: 397～408

9 Shevchenko A, Jensen ON, Podtelejnikov AV *et al*. Linking genome and proteome by mass spectrometry: large-scale identification of yeast proteins from two dimensional gels. Proc Natl Acad Sci USA, 1996, 93: 14440～14445

10 Gras R, Muller M, Gasteiger E *et al*. Improving protein identification from peptide mass fingerprinting through a parameterized multi-level scoring algorithm and an optimized peak detection. Electrophoresis, 1999, 20: 3535～3550

11 Hoogland C, Sanchez JC, Tonella L *et al*. The 1999 SWISS-2DPAGE database update. Nucleic Acids Res, 2000, 28: 286～288

12 Rabilloud T, Kieffer S, Procaccio V *et al*. Two-dimensional electrophoresis of human placental mitochondria and protein identification by mass spectrometry: toward a human mitochondrial proteome. Electrophoresis, 1998, 19: 1006～1014

13 Link AJ, Eng J, Schieltz DM *et al*. Direct analysis of protein complexes using mass spectrometry. Nat Biotechnol, 1999, 17: 676～682

14 Jung E, Heller M, Sanchez JC *et al*. Proteomics meets cell biology: the establishment of subcellular proteomes. Electrophoresis, 2000, 21: 3369～3377

15 Alberts B, Miake-Lye R. Unscrambling the puzzle of biological machines: the importance of the details. Cell, 1992, 722: 203～233

16 Bartlett SE, Hendry IA. Identifying the G protein, Gz alpha, and its associated proteins in nervous tissue using mass spectrometry and microsequencing techniques. Int J Dev Neurosci, 1997, 15: 267～274

17 Heller M, Goodlett DR, Watts JD *et al*. A comprehensive characterization of the T-cell antigen receptor complex compo-

sition by microcapillary liquid chromatography-tandem mass spectrometry. Electrophoresis, 2000, 21:2180～2195
18 Pandey A, Podtelejnikov AV, Blagoev B *et al*. Analysis of receptor signaling pathways by mass spectrometry: identification of vav-2 as a substrate of the epidermal and platelet-derived growth factor receptors. Proc Natl Acad Sci USA, 2000, 97:179～184
19 Lopez MF, Kristal BS, Chernokalskaya E *et al*. High-throughput profiling of the mitochondrial proteome using affinity fractionation and automation. Electrophoresis, 2000, 21:3427～3440
20 Fialka I, Steinlein P, Ahorn H *et al*. Identification of syntenin as a protein of the apical early endocytic compartment in Madin-Darby canine kidney cells. J Biol Chem, 1999, 274:26233～26239
21 Rout MP, Aitchison JD, Suprato A *et al*. The yeast nuclear pore complex: composition, architecture, and transport mechanism. J Cell Biol, 2000, 148:635～652
22 Wigge PA, Jensen ON, Holmes S *et al*. Analysis of the Saccharomyces spindle pole by matrix-assisted laser desorption/ionization (MALDI) mass spectrometry. J Cell Biol, 1998, 141:967～977
23 Morre DJ, Howell KE, Cook GMW *et al*. Cell free analysis of membrane traffic. Alan R. Liss, New York: 1988. 317～331
24 Spector DL, Golgman RD, Leinward LA eds. Cells a laboratory mannual. 1998.
25 Dealtry GB, Rickwood D eds. Cell biology labfax. London: Scientific Publishers, 1992
26 Rickwood D, Ford TC, Graham JM. Nycodenz: a new nonionic iodinated gradient medium. Anal Biochem, 1982, 123: 23～31
27 Rickwood D ed. Iodinated density grsdient media-A practical approach, IRL Press at Oxford University Press, United Kingdom. 1983. 1
28 Wattiaux R, Wattiaux-De Coninck S. Separation of Organelles. In(Rickwood D. ed.) Iodinated density grsdient media-A practical approach. London: IRL Press, 1983. 119
29 Rout MP, Field MC. Isolation and characterization of subnuclear compartments from Trypanosoma brucei. Identification of a major repetitive nuclear lamina component. J Biol Chem, 2001, 276:38261～38271
30 Cahill A, Baio DL, Cunningham CC. Isolation and characterization of rat liver mitochondrial ribosomes. Anal Biochem, 1995, 232:47～55
31 Murayama K, Fujimura T, Morita M *et al*. One-step subcellular fractionation of rat liver tissue using a Nycodenz density gradient prepared by freezing-thawing and two-dimensional sodium dodecyl sulfate electrophoresis profiles of the main fraction of organelles. Electrophoresis, 2001, 22:2872～2880
32 Willkins MR, Pasquali C, Appel RD *et al*. From proteins to proteomes: large scale protein identification by two-dimensional electrophoresis and amino acid analysis. Biotechnology, 1996, 14:61～65
33 Packer NH, Lawson MA, Jardine DR *et al*. Analyzing glycoproteins separated by two-dimensional gel electrophoresis. Electrophoresis, 1998, 19:981～988
34 Packer NH, Willkins MR, Golaz O *et al*. Characterization of human plasma glycoproteins separated by two-dimensional gel electrophoresis. Biotechnology, 1996, 14:66～70
35 Eng JK, McCormack AL, Yates JR. An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. J Amer Soc Mass Spectrom, 1994, 5:976～989
36 Link AJ, Carmack E, Yates JR. A strategy for the identification of proteins localized to subcellular spaces: applications to E. coli periplasmic proteins. Int J Mass Spectrom Ion Proc, 1997, 160:303～316
37 McCormack AL, Schieltz DM, Goode B *et al*. Direct analysis and identification of proteins in mixtures by LC/MS/MS and database searching at the low-femtomole level. Anal Chem, 1997, 69:767～776
38 Wolters DA, Washburn MP, Yates Ⅲ JR. An automated multidimensional protein identification technology for shotgun proteomics. Anal Chem, 2001, 73:5683～5690
39 Rigaut G, Shevchenko A, Rutz B *et al*. A generic protein purification method for protein complex characterization and proteome exploration. Nat Biotechnol, 1999, 17:1030～1032
40 Murphy RF, Boland MV, Velliste M *et al*. Towards a systematics for protein subcelluar location: quantitative descrip-

tion of protein localization patterns and automated analysis of fluorescence microscope images. Proc Int Conf Intell Syst Mol Biol,2000,8:251～9

41 Neubauer G,King A,Rappsilber J *et al*. Mass spectrometry and EST-database searching allows characterization of the multi-protein spliceosome complex. Nat Genet,1998,20:46～50

42 Nakai K. Protein sorting signals and prediction of subcellular localization. Adv Protein Chem,2000,54:277～344

43 Nakai K. Review: Prediction of in vivo fates of proteins in the era of genomics and proteomics. J Struct Biol,2001, 134:103～116

44 Emanuelsson O,Heijne GV. Prediction of organellar targeting signals. Biochim Biophys Acta,2001,1541:114～119

45 Feng ZP. An overview on predicting the subcellular location of a protein. In silico Biology,2002,2:291～303

46 Emanuelssen O,Nielsen H,Heijne GV *et al*. A neural network based method for predicting xhloroplast transit peptides and their cleavage sites. Protein Sci,1999, 8:978～984

47 Claros M. MitoProt,a macintosh application for studying mitochondrial proteins. Comput Appl Biosci,1995,11:441～447

48 Claros M,Vincens P *et al*. Computational method to predict mitochondrially imported proteins and their targeting sequences. Eur J Biochem,1996,241:779～786

49 Nielsen H,Engelbrecht J,Brunak S *et al*. Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. Protein Eng,1997,10:1～6

50 Cokol M,Nair R,Rost B. Finding nuclear localization signals. EMBO Rep,2000,1:411～415

51 Gierasch LM. Signal sequences. Biochemistry,1989,28:923～930

52 Krogh A,Larsson B,von Heijne G *et al*. Predicting transmembrane protein topology with a hid-den Markov model: Application to complete genomes. J Mol Biol,2001,305: 567～580

53 Tusna´dy GE,Simon I. Principles governing amino acid composition of integral membrane proteins: Application to topology prediction. J Mol Biol,1998,283:489～506

54 Hua S,Sun Z. Support vector machine approach for protein subcellular localization prediction. Bioinformatics,2001,17: 721～728

55 Reinhardt A,Hubbard T. Using neural networks for prediction of the subcellular location of proteins. Nucleic Acids Res,1998,26:2230～2236

56 Emanuelsson O,Nielsen H,Brunak S *et al*. Prediction subcellular localization of proteins based on their N-terminal amino acid sequence. J Mol Biol,2000,300: 1005～1016

57 Nakai K,Kanehisa M. Expert system for predicting protein localization sites in Gram-negative bacteria Proteins,1991, 11:95～110

58 Nakai K,Kamehisa M. A knowledge base for predicting protein localization sites in eukaryotic cells. Genomics,1992, 14:897～911

59 Horton P,Nakai K. A probabilistic classification system for predicting the cellular localization sites of proteins. Proc Int Cont Intellig Syst Mol Biol,1996,4:109～115

60 Horton P,Nakai K. Better prediction of protein cellular localization sites with the k nearest neighbor classifer. Proc Int Cont Intellig Syst Mol Biol,1997,5: 147～152

61 Nakai K,Horton P. PSORT: a program for detecting the sorting signals of proteins and predicting their subcellular localization. Trends Biochem Sci,1999,24:34～36

62 Nakashima H,Nishikawa K. Discrimination of intracellular and extracellular proteins using amino acid composition and residue-pair frequencies. J Mol Biol,1994,238:54～61

63 Cedano J,Aloy P,Perez-Pons JA *et al*. Relation between amino acid composition and cellular location of proteins. J Mol Biol,1997,266:594～600

64 Chou KC,Elrod DW. Using discriminant function for prediction of subcellular location of prokaryotic proteins. Biochem Biophys Res Commun,1998,252:63～68

65 Chou KC,Elrod DW,Protein subcellular location prediction. Protein Eng,1999,12: 107～118

66 Chou KC, Elrod DW, Prediction of membrane protein types and subcellular locations. Proteins, 1999, 34: 137～153
67 Yuan Z. Prediction of protein subcellular locations using Markov chain models. FEBS Lett, 1999, 451: 23～26
68 Stanford DR, Martin NC, Hopper AK. ADEPTs: Information necessary for subcellular distribution of eukaryotic sorting isozymes resides in domains missing from eubacterial and archaeal counterparts. Nucleic Acids Res, 2000, 28: 383～392
69 Marcotte EM, Xenarios I, Van Der Bliek AM *et al*. Localizing proteins in the cell from their phylogenetic profiles. Proc Natl Acad Sci USA, 2000, 97: 12115～12120
70 Jung E, Hoogland C, Chiappe D *et al*. The establishment of a human liver nuclei two-dimensional electrophoresis reference map. Electrophoresis, 2000, 21: 3483～3487
71 Chataway TK, Whittle AM, Lewis MD *et al*. Two-dimensional mapping and microsequencing of lysosomal proteins from human placenta. Placenta, 1998, 19: 643～654
72 Rabilloud T, Kieffer S, Procaccio V *et al*. Two-dimensional electrophoresis of human placental mitochondria and protein identification by mass spectrometry: toward a human mitochondrial proteome. Electtrophoresis, 1998, 19: 1006～1014
73 Millar AH, Sweetlove LJ, Giege P. Analysis of the Arabidopsis mitochondrial proteome. Plant Physiol, 2001, 127: 1711～1727
74 Scherl A, Coute Y, Deon C *et al*. Functional proteomic analysis of human nucleolus. Mol Biol Cell, 2002, 13: 4100～4109
75 Sanchez JC, Chiappe D, Converset V *et al*. The mouse SWISS-2DPAGE database: a tool for proteomics study of diabetes and obesity. Proteomics, 2001, 1: 136～163
76 Muller EC, Schumann M, Rickers A *et al*. Study of Burkitt lymphoma cell line proteins by high resolution two-dimensional gel electrophoresis and nanoelectrospray mass spectrometry. Electrophoresis, 1999, 20: 320～330
77 Gerner C, Holzmann K, Grimm R *et al*. Similarity between nuclear matrix proteins of various cells revealed by an improved isolation method. J Cell Biochem, 1998, 71: 363～374.
78 Gerner C, Sauermann G. Nuclear matrix proteins specific for subtypes of human hematopoietic cells. J Cell Biochem, 1999, 72: 470～482
79 Gerner C, Holzmann K, Meissner M *et al*. Reassembling proteins and chaperones in human nuclear matrix protein fractions. J Cell Biochem, 1999, 74: 145～151
80 Link AJ, Eng J, Schieltz DM *et al*. Direct analysis of protein complexes using mass spectrometry. Nat Biotechnol, 1999, 17: 676～682.
81 Yoon HJ, Feoktistova A, Wolfe BA *et al*. Proteomics analysis identifies new components of the fission and budding yeast anaphase-promoting complexes. Curr Biol, 2002, 12: 2048～2054
82 Ohi MD, Link AJ, Ren L *et al*. Proteomics analysis reveals stable multiprotein complexes in both fission and budding yeasts containing Myb-related Cdc5p/Cef1p, novel pre-mRNA splicing factors, and snRNAs. Mol Cell Biol, 2002, 22: 2011～2024
83 Fountoulakis M, Berndt P, Langen H *et al*. The rat liver mitochondrial proteins. Electrophoresis, 2002, 23: 311～328
84 Alberts B *et al*. Molecular Biology of the Cell. 3th ed. New York and London: Garland publishing InC. 1994, 477～647
85 Darnell J *et al*. Molecular Cell Biology. 3th ed. New York: Scientific American Books. 1995. 569～612, 617～663
86 Adessi C, Miege C, Albrieux C *et al*. Two-dimensional electrophoresis of membrane proteins: a current challenge for immobilized pH gradients. Electrophoresis, 1997, 18: 127～135
87 Santoni V, Molloy M, Rabilloud T *et al*. Membrane proteins and proteomics: un amour impossible? Electrophoresis, 2000, 21: 1054～1070
88 Wilkins MR, Gasteiger E, Sanchez JC *et al*. Two-dimensional gel electrophoresis for proteome projects: the effects of protein hydrophobicity and copy number. Electrophoresis, 1998, 19: 1501～1505
89 Rimpliainen MA, Righetti PG. Membrane protein analysis by isoelectric focusing in immobilized pH gradients. Electrophoresis, 1985, 6: 419～422
90 Righetti PG, Gelfi C, Bossi ML *et al*. Isoelectric focusing and non-isoelectric precipitation of ferritin in immobilized pH

gradients: an improved protocol overcoming protein-matrix interactions. Electrophoresis, 1987, 8:62～70

91 Molloy MP. Two-dimensional electrophoresis of membrane proteins using immobilized pH gradients. Anal Biochem, 2000, 280:1～10

92 Rabilloud T, Adessi C, Giraudel A *et al*. Improvement of the solubilization of proteins in two-dimensional electrophoresis with immobilized pH gradients. Electrophoresis, 1997, 18:307～316

93 Perdew GH, Schaup HW, Selivonchick DP. The use of a zwitterionic detergent in two-dimensional gel electrophoresis of trout liver microsomes. Anal Biochem, 1983, 135:453～455

94 Hochstrasser DF, Harrington MG, Hochstrasser AC *et al*. Methods for increasing the resolution of two-dimensional protein electrophoresis. Anal Biochem, 1988, 173: 424～435

95 Holloway PJ, Arundel PH. High-resolution two-dimensional electrophoresis of plant proteins. Anal Biochem, 1988, 172: 8～15

96 Ruegg UT, Rudinger J. Reductive cleavage of cystine disulfides with tributylphosphine. Methods Enzymol, 1977, 48: 111～116

97 Herbert BR, Molloy MP, Gooley AA *et al*. Improved protein solubility in two-dimensional electrophoresis using tributyl phosphine as reducing agent. Electrophoresis, 1998, 19:845～851

98 Molloy MP, Herbert BR, Walsh BJ *et al*. Extraction of membrane proteins by differential solubilization for separation using two-dimensional gel electrophoresis. Electrophoresis, 1998, 19: 837～844

99 Bordier C. Phase separation of integral membrane proteins in Triton X-114. J Biol Chem, 1981, 256:1604～1607

100 Fujiki Y, Hubbard AL, Fowler S *et al*. Isolation of intracellular membranes by means of sodium carbonate treatment. J Cell Biol, 1982, 93:97～102

101 Molloy MP, Herbert BR, Williams KL *et al*. Extraction of Escherichia coli proteins with organic solvents prior to two-dimensional electrophoresis. Electrophoresis, 1999, 20:701～704

102 Molloy MP, Herbert BR, Slade MB *et al*. Proteomic analysis of the Escherichia coli outer membrane. Eur J Biochem, 2000, 267:2871～2881

103 Santoni V, Rabilloud T, Doumas P *et al*. Towards the recovery of hydrophobic proteins on two-dimensional electrophoresis gels. Electrophoresis, 1999, 20:705～711

104 Santoni V, Kieffer S, Desclaux D *et al*. Membrane proteomics: Use of additive main effects with multiplicative interaction model to classify plasma membrane proteins according to their solubility and electrophoretic properties. Electrophoresis, 2000, 21:3329～3344

105 Galevav, AltermannM. Comparison of one-dimensional and two-dimensional gel electrophoresis as a separation tool for proteomic analysis of rat liver microsomes: cytochromes P450 and other membrane proteins. Proteomics, 2002, 2: 713～722

106 Washburn MP, Wolters D, Yates Ⅲ JR. Large-scale analysis of the yeast proteome by multidimensional protein identification technology. Nat Biotechnol, 2001, 19:242～247

107 Alberts B *et al*. Molecular Biology of the Cell. 3th ed. New York and London: Garland publishing InC. 1994. 551～617

108 Darnell J *et al*. Molecular Cell Biology. 3th ed. New York: Scientific American Books. 1995

109 Dunphy W, Rothman JE. Compartmental organization of the Golgi stack. Cell, 1985, 42:13～21

110 Mellman I, Simons K. The Golgi complex: in vitro veritas? Cell, 1992, 68:829～840

111 Taylor RS, Wu CC, Hays LG *et al*. Proteomics of rat liver Golgi complex: Minor proteins are identified through sequential fractionation. Electrophoresis, 2000, 21:3441～3459

112 Howell KE, Palade GE. Hepatic Golgi fractions resolved into membrane and content subfractions. J Cell Biol, 1982, 92:822～832

113 Taylor RS, Jones SM, Dahl RH *et al*. Characterization of the Golgi complex cleared of proteins in transit and examination of calcium uptake activities. Mol Biol Cell, 1997, 8:1911～1931

114 Taylor RS, Fialka I, Jones SM *et al*. Two-dimensional mapping of the endogenous proteins of the rat hepatocyte Golgi complex cleared of proteins in transit. Electrophoresis, 1997, 18:2601～2612

115 Wu CC, Taylor RS, Lane DR *et al*. GMx33: A novel family of trans-Golgi proteins identified by proteomics. Traffic, 2000, 1:963～975
116 Bell AW, Ward MA, Blackstock WP *et al*. Proteomics characterization of abundant Golgi membrane proteins. J Biol Cell, 2001, 276:5152～5165
117 Clermont Y, Xia L, Rambourg A *et al*. Transport of casein submicelles and formation of secretion granules in the Golgi apparatus of epithelial cells of the lactating mammary gland of the rat. Anat Rec, 1993, 235:363～373
118 Wooding FB. Comparative mammary fine structure. In Peaker M ed. Comparative Aspects of lactation. London: Academic Press, 1977. 1～41
119 Baldwin RL, Yang YT. Enzymatic and metabolic changes in the development of lactation. In: Larson BL, Smith VR ed. New York and London: Academic Press, 1974. 349～411
120 Wu CC, Yates Ⅲ JR, Neville MC *et al*. Proteomic analysis of two functional states of the Golgi complex in mammary epithelial cells. Traffic, 2000, 1:769～782
121 Alberts B *et al*. Molecular Biology of the Cell. 3th ed. New York and London: Garland publishing InC. 1994. 335～399
122 Ris H. The 3-D structure of the nuclear pore complex as seen by high voltage electron microscopy and high resolution low voltage scanning electron microscopy. EMSA Bull, 1991, 21:54～56
123 Akey CW. The nuclear pore complex. Curr Opin Struct Biol, 1992, 2:258～263
124 Davis LI, The nuclear pore complex. Annu Rev Biochem, 1995, 64:865～896
125 Fabre E, Hurt E. Yeast genetics to dissect the nuclear pore complex and nucleocytoplasmic trafficking. Annu Rev Genet, 1997, 31:277～313
126 Lain S, Xirodimas D, Lane DP. Accumulating active p53 in the nucleus by inhibition of nuclear export: a novel strategy to promote the p53 tumor suppressor function. Exp Cell Res, 1999, 253:315～324
127 Jeffries S, Capobianco AJ. Neoplastic transformation by Nortch requires nuclear localization. Mol Cell Biol, 2000, 20:3928～3941
128 Takizawa CG, Morgan DO. Control of mitosis by changes in the subcellular location of cyclin-B1-Cdk1 and Cdc25C. Curr Opin Cell Biol, 2000, 12:658～665
129 Buss F, Kent H, Stewart M *et al*. Role of different domains in the self-association of rat nucleoporin p62. J Cell Sci, 1994, 107:631～638
130 Radu A, Moore MS, Blobel G. The peptide repeat domain of nucleoporin Nup98 functions as a docking site in transport across the nuclear pore complex. Cell, 1995, 81:215～222
131 Kehlenbach RH, Dickmanns A, Kehlenbach A *et al*. A role for RanBP1 in the release of CRM1 from the nuclear pore complex in a terminal step of nuclear export. J. Cell Biol, 1999, 145:645～657
132 Mattaj IW, Englmeier L. Nucleocytoplasmic transport: the soluble phase. Annu Rev Biochem, 1998, 67:265～306
133 Talcott B, Moore MS. Getting across the nuclear pore complex. Trends Cell Biol, 1999, 9:312～318
134 Rout MP, Blobel G. Isolation of the yeast nuclear pore complex. J Cell Biol, 1993, 123:771～783
135 Rout MP, Wente SR. Pores for thought: nuclear pore proteins. Trends Cell Biol, 1994, 4:357～365
136 Yang Q, Rout MP, Akey CW. Three-dimensional architecture of the isolated yeast nuclear pore complex: functional and evolutionary implications. Mol Cell, 1998, 1:223～234
137 Cronshaw JM, Krutchinsky AN, Zhang W *et al*. Proteomics analysis of the mammalian nuclear pore complex. J Cell Biol, 2002, 158:915～927
138 Ryan KJ, Wente SR. The nuclear core complex: a protein machine bridging the nucleus and cytoplasm. Curr Opin Cell Biol, 2000, 12:361～371
139 Akey CW, Radermacher M. Architecture of the Xenopus nuclear pore complex revealed by three-dimensional cryo-electron microscopy. J Cell Biol, 1993, 122:1～19
140 Reichelt R, Holzenburg A, Buhle EL *et al*. Correlation between structure and mass distribution of the nuclear pore complex and of distinct pore complex components. J Cell Biol, 1990, 110:883～894

141 Blobel G, Wozniak RW. Proteomics for the pore. Nature, 2000, 403:835～836
142 Shaw PJ, Jordan EG. The nucleolus. Annu Rev Cell Dev Biol, 1995, 11:93～121
143 Scheer U, Hock R. Structure and function of the nucleolus. Curr Opin Cell Biol, 1999, 11:385～390
144 Venema J, Tollevery D. Ribosome synthesis in *Saccharomyces cerevisiae*. Annu Rev Genet, 1999, 33:261～311
145 Carmo-Fonseca M, Mendes-Soares L, Campos I. To be or not to be in the nucleolus. Nat Cell Biol, 2000, 2: E107～112
146 Dundr M, Misteli T, Olson MO. The dynamics of postmitotic reassembly of the nucleolus. J Cell Biol, 2000, 150: 433～446
147 Pederson T. The plurifunctional nucleolus. Nucleic Acids Res, 1998, 26:3871～3876
148 Olson MO, Dundr M, Szebeni A. The nucleolus: an old factory with unexpected capabilities. Trends Cell Biol, 2000, 10:189～196
149 Politz JC, Yarovoi S, Kilroy SM *et al*. Signal recognition particle components in the nucleolus. Proc Natl Acad Sci USA, 2000, 97:55～60
150 Lange TS, Gerbi SA. Transient nucleolar localization of U6 small nuclear RNA in Xenopus Laevis oocytes. Mol Biol Cell, 2000, 11:2419～2428
151 Mitchell JR, Wood E, Collins K. A telomerase component is defective in the human disease dyskeratosis congenita. Nature, 1999, 402:551～555
152 Schneiter R, Kadowaki T, Tartakoff AM. mRNA transport in yeast: time to reinvestigate the functions of the nucleolus. Mol Biol Cell, 1995, 6:357～370
153 Bertrand E, Houser-Scott F, Kendall A *et al*. Nucleolar localization of early tRNA processing. Genes Dev, 1998, 12: 2463～2468
154 Visintin R, Amon A. The nucleolus: the magician's hat for cell cycle tricks. Curr Opin Cell Biol, 2000, 12:752
155 Alberts B *et al*. Molecular Biology of the Cell. 3rd ed. New York and London: Garland publishing InC. 1994. 401～474
156 Andersen JS, Lyon CE, Fox AH *et al*. Directed proteomic analysis of the human nucleolus. Curr Biol, 2002, 12:1～11
157 Ghosh S. The nucleolar structure. Int Rev Cytol, 1976, 44:1～28
158 Fox AH, Lam YW, Leung A *et al*. Paraspeckles: a novel nuclear domain. Curr Biol, 2002, 12:13～25
159 Pederson T. Proteomics of the nucleolus: more proteins, more functions? Trends Biochem Sci, 2002, 27:111～112
160 Schuldt A. Proteomics of the nucleolus. Nat Cell Biol, 2002. 4: E35
161 Dundr M, Misteli T. Nucleomics: an inventory of the nucleolus. Mol Cell, 2002, 9:5～7
162 Ospina JK, Matera AG. Proteomics: the nucleolus weighs in. Curr Biol, 2002, 12: R29～R31
163 Kiberstis PA, Mitochondria make a comeback. Science, 1999, 283:1475
164 Alberts B *et al*. Molecular Biology of the Cell. 3rd ed. New York and London: Garland publishing InC. 1994. 653～720
165 Gunter TE, Gunter KK, Sheu SS *et al*. Mitochondrial calcium transport: physiological and pathological relevance. Am J Physiol, 1994, 267:C313～C339
166 Hajnoczky G, Robb-Gaspers LD, Seitz MB *et al*. Decoding of cytosolic calcium oscillations in the mitochondria. Cell, 1995, 82:415～424
167 Rizzuto R, Simpson AWM, Brini M *et al*. Rapid changes of mitochondrial Ca^{2+} revealed by specifically targeted recombinant aequorin. Nature, 1992, 358:325～327
168 Earnshaw WC. Apoptosis. A cellular poison cupboard. Nature, 1999, 397:387～389
169 Green DR, Reed JC. Mitochondria and apoptosis. Science, 1998, 281:1309～1312
170 Follstad BD, Wang DI. Stephanopoulos G, Mitochondrial membrane potential differentiates cells resistant to apoptosis in hybridoma cultures. Eur J Biochem, 2000, 267:6534～6540
171 Kitagawa A. Effects of cresols (o-, m-, and p-isomers) on the bioenergetic system in isolated rat liver mitochondria. Drug Chem Toxicol, 2001. 24:39～47

172 de Duve C. Tissue fractionation—Past and present. J Cell Biol, 1971, 50:20～55

173 de Duve C, Beaufay H. A short history tissue fractionation. J Cell Biol, 1981, 91: 293～299

174 Graham J, Ford T, Rickwood D. The preparation of subcellular organelles from mouse liver in self-generated gradients of Iodixanol. Anal Biochem, 1994, 20:367～373

175 Kruft V, Euble H, Jansch L *et al*. Proteomic approach to identify novel mitochondrial proteins in Arabidopsis. Plant Physiol, 2001, 127:1694～1710

176 Millar AH, Sweetlove LJ, Giege P. Analysis of the Arabidopsis mitochondrial proteome. Plant Physiol, 2001, 127: 1711～1727

177 Hanson BJ, Schulenberg B, Patton WF. A novel subfractionation approach for mitochondrial proteins: a three-dimensional mitochondrial proteome map. Electrophoresis, 2001, 22:950～959

178 Werhahn W, Braun HP. Biochemical dissection of the mitochondrial proteome from Arabidopsis thaliana by three-dimensional gel electrophoresis. Electrophoresis, 2002, 23:640～646

179 Jansch L, Kruft V, Schmitz UK. Unique composition of the preprotein translocase of the outer mitochondrial membrane from plants. J Biol Chem, 1998, 273:17251～17257

180 Werhahn W, Niemeyer A. Jansch L. Purification and characterization of the preprotein translocase of the outer mitochondrial membrane from Arabidopsis. Identification of multiple forms of TOM20. Plant Physiol, 2001, 125:943～954

181 O'Brien TW. The general occurrence of 55 S ribosomes in mammalian liver mitochondria. J Biol Chem, 1971, 246: 3409～3417

182 Saachi A, Cerbone F, Cammarrano P *et al*. Physiochemical characterization of ribosome-like (55-S) particles from rat liver mitochondria. Biochim Biophys Acta, 1973, 308:390～403

183 Hamilton MG, O'Brien TW. Ultracentrifugal characterization of the mitochondrial ribosome and subribosomal particles of bovine liver: molecular size and composition. Biochemistry, 1974, 13:5400～5403

184 Koc EC, Burkhart W, Blackburn K *et al*. A proteomics approach to the identification of mammalian mitochondrial small subunit ribosome proteins. J Biol Chem, 2000, 275:32585～32591

185 Koc EC, Burkhart W, Blackburn K *et al*. Identification of four proteins from the small subunit of the mammalian mitochondrial ribosome using a proteomics approach. Protein Sci, 2001, 10:471～481

186 Suzuki T, Terasaki M, Takemoto-Hori C *et al*. Proteomic analysis of the mammalian mitochondrial ribosome——identification of protein components in the 28S small subunit. J Biol Chem, 2001, 276:33181～33195

187 Suzuki T, Terasaki M, Takemoto-Hori C *et al*. Structural compensation for the deficit of rRNA with proteins in the mammalian mitochondrial ribosome——systematic analysis of protein components of the large ribosomal subunit from mammalian mitochondria. J Biol Chem, 2001, 276:21724～21736

188 Kim SH, Vlkolinsky R, Cairns N *et al*. Decreased levels of complex Ⅲ core protein 1 and complex Ⅴ beta chain in brains from patients with Alzheimer's diseases and Down syndrome. Cell Mol Life Sci, 2000, 57:1810～1816

189 Kim SH, Fountoulakis M, Dierssen M *et al*. Decreased protein levels of complex Ⅰ 30-Kda subunit in fetal Down syndrome brains. J Neural Transm-Supp, 2001, 61:109～116

190 Tonge RP, Shaw J, Wallis N *et al*. Toxicol Lett, 2000, 116(suppl): 114

191 Ruepp SU, Tonge RP, Shaw J *et al*. Genomics and proteomics analysis of acetaminophen toxicity in mouse liver. Toxicol Sci, 2002, 65:135～150

192 Krapfenbauer K, Berger M, Lubec G *et al*. Changes in the brain protein level following administration of kainic acid. Electrophoresis, 2001, 22:2086～2091

193 Bruce SK, Abraham, MB, Apoptogenic ganglioside GD3 directly induces the mitochondrial permeability transition. J Biol Chem, 1999, 274:23169～23175

194 Yan JX, Wait R, Berkelman T *et al*. A modified silver staining protocol for visualization of proteins compatible with matrix-assisted laser desorption/ionization and electrospray ionization-mass spectrometry. Electrophoresis, 2000, 21: 3666～3672

第二十章　蛋白质组研究中的非凝胶技术

第一节　概　　况

在蛋白质组学的研究技术中，双向凝胶电泳（2D-PAGE）由于具有无可比拟的高分辨率而始终占据着重要的地位。然而这种技术本身也存在一些难以克服的缺陷，如在双向凝胶电泳图谱上，并非每个蛋白质点所包含的蛋白质量都足以用于鉴定蛋白质；操作费时费力，难以实现和质谱的直接联用，不易自动化；存在偏向性问题，如对极大蛋白质（相对分子质量＞200 000）、极小蛋白质（相对分子质量＜8 000）、低丰度蛋白质（如信号蛋白和转录因子）、极碱性蛋白质和疏水性蛋白质（膜结合蛋白质和跨膜蛋白质），还难以进行分离分析。这些缺陷都在一定程度限制了双向凝胶电泳的应用。因此，发展新型高分辨率、高通量、高峰容量的分离分析方法已经成为加速蛋白质组学研究的关键。

近年来，分离科学发展出了一些新型的分离模式，使高效液相色谱、毛细管电泳和毛细管电色谱等技术在蛋白质组学研究中焕发出了新的活力。这些新技术，有的是不以凝胶为分离介质的，我们暂且将其归为蛋白质组学研究中的非凝胶技术。这些非凝胶分离技术所表现出的高效、快速、自动化程度高、高灵敏度、低检测限等优点已经极大并有可能进一步推动蛋白质组学的相关研究。

对于采用传统双向凝胶电泳技术进行蛋白质组学的研究步骤，本书第十三章到十七

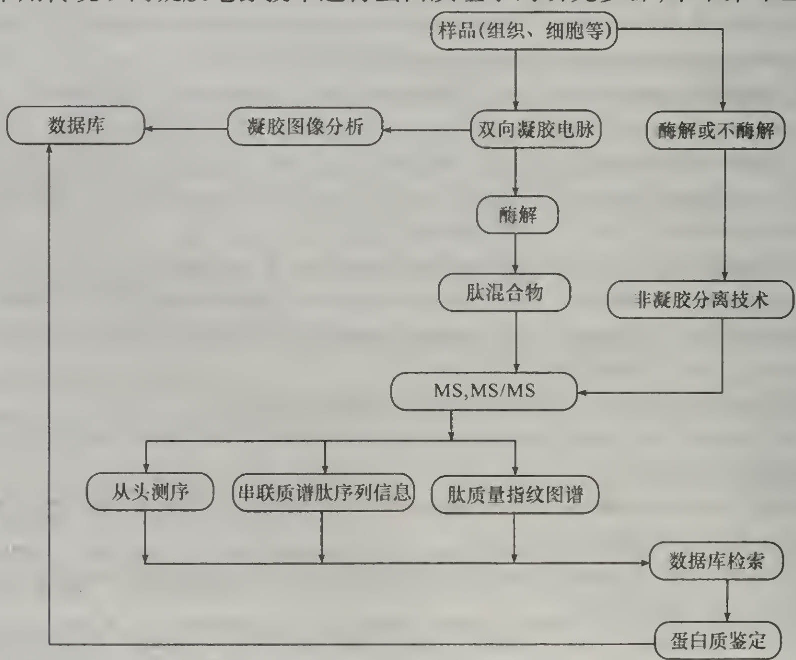


图 20.1　凝胶和非凝胶技术研究蛋白质组学的技术路线

章已进行了较为详细的介绍和探讨;而采用非凝胶分离技术进行蛋白质组学研究,其基本步骤又有何不同呢? 两者之间的对比如图 20.1 所示。

可以看出,非凝胶分离技术的介入大大简化了蛋白质组学的研究步骤。在此将根据非凝胶技术所基于的分离模式,分别从非凝胶色谱、非凝胶电泳(液相等电聚焦和毛细管电泳)几方面对非凝胶分离技术在蛋白质组学中的应用作一阐述,最后对新兴的微分离技术——毛细管电色谱进行简要介绍。

第二节　非凝胶色谱技术

一、一维色谱及其与质谱联用技术

如前所述,高效液相色谱经过几十年的发展,已经派生出多种分离模式。但由于在蛋白质组学的研究当中,最常用的质谱离子源是电喷雾离子源(ESI)和基质辅助激光解吸离子源(MALDI),这两种离子源都要求整个质谱系统中最好不被引入过高浓度的盐。所以,在将一维色谱和质谱联用时,一般是将体系中没有盐存在的反相高效液相色谱(RP-HPLC)和质谱系统联用,从而达到最终的分离鉴定目的。

比较方便而简单的联用技术是直接在电喷雾质谱里喷雾针所用的毛细管内填充一段反相液相色谱的填料,实现分离和喷雾一步完成。在某种程度上,这种技术相当于将高效液相色谱"整合"到质谱系统中。Davis 等人[1]将反相 C18 填料填充到 15cm×250μm i. d. 的毛细管喷针内,前端以 PVDF 膜作为塞板将填料加以固定,乙腈梯度洗脱肌血球素和细胞色素 C 的酶解产物,串联质谱检测。结果表明,采用这种技术,多肽的最低检测限达到了 10 fmol 级。但是将 PVDF 膜和反相填料先后填充在质谱的毛细管喷针当中需要较高的操作技巧,这在一定程度上限制了这种方法的发展。后来,Moore 等人[2]采用原位聚合(*in situ* polymerization)技术在电喷雾的喷针内就地聚合了一段苯乙烯-二乙烯基苯的多孔高分子聚合物作为反相分离介质,串联质谱检测和鉴定了标准蛋白质,然后比较了这种原位柱和 C18 填充柱对四种标准蛋白质的分离能力,结果发现前者具有更高的分离效率,这种技术大大简化了整个一维色谱-质谱体系的建立过程。在此基础上,Premstaller 等人[3]又优化了苯乙烯-二乙烯基苯的聚合条件,得到了更高效的反相多孔聚合物原位柱,并将这种原位柱和 ESI-MS 联用,对蛋白质及牛过氧化氢酶的胰蛋白酶解产物进行了很好的分离鉴定,最高检测限达到了飞摩尔,准确度可达 50～150 ppm,使原位多孔聚合物材料在蛋白质组学研究中得到更好的应用。目前,内部装有填料的喷针已经实现了商品化。

与此同时,相对于"整合"技术,一些研究小组也在开展相对独立的 HPLC 和 MS 的直接联用技术,解决了一些原先采用双向凝胶电泳难以解决的问题,并将一维色谱与质谱的联用技术应用到更广阔的范围。例如,大脑神经肽的相对分子质量一般为 300～5 000,用传统双向凝胶电泳分离存在很大困难,而 Skodld 等[4]以纳升级液相色谱-电喷雾-时间飞行-质谱(nanoliter-LC-ESI-Q-TOF-MS)联用技术成功分离并鉴定了小鼠大脑某个确定区域的 1 500 个肽段,最后,他们认为这种联用技术是研究大脑中大量肽段和蛋白质的强有力手段。Blonder 等人[5]通过改进对膜蛋白的处理技术,成功地用毛细管液相色谱-串联

质谱(capillary LC-iontrap-MS/MS)技术对从 *D. radiodurans* 分离的膜蛋白配料混合物进行了分离和鉴定。Spahr 等用[6]LC-MS/MS (quadrupole-time of flight MS, Q-TOF)对商业收集的男性尿样进行了 4 次分析,成功鉴定了 124 种蛋白质。McComack 等人[7]将 8 种蛋白质的蛋白酶酶解产物进样到一个内填 C 18 反相色谱填料、内径为 75 μm 的毛细管中,乙腈梯度洗脱肽段的混合物,电喷雾串联质谱(nanoscale HPLC-ESI MS/MS)分离鉴定了数以千计的多肽。Zhou 等人[8]利用 LC-MS/MS 对两种蛋白质定量方法——同位素编码亲和标签技术和稳定同位素标记技术,进行了比较和评估,认为后者更简单、高效和灵敏,使一维色谱与质谱联用技术成为蛋白质定量方法的评估手段之一。

在一维色谱-质谱联用技术中得到具体应用的同时,从方法学的角度来说,其分离和质谱鉴定条件的优化也就成为另一个研究热点。Vissers 等[9]提出了一种流动相流速可变的“峰肼”色谱接口技术,可以保证分离样品不需经过分流技术而得到分离鉴定。Shen 等人[10]详细讨论了毛细管液相色谱-质谱体系中毛细管内径、长度、多孔与非多孔填料以及流速等对峰容量的影响,发现降低 LC 流动相的流速对低丰度蛋白质的检测具有提高灵敏度的效果;在回收率方面,他们发现使用多孔或无孔填料对多肽的质量回收率几乎没有影响,表明在这一联用系统中,可以尽可能选择具有高分离效率的多孔填料。Laurell 等[11]则将固定化酶技术引入 LC-MS 系统,使蛋白质分析的自动化程度得到提高。这些研究对一维色谱-质谱联用技术在方法学上都具有十分重要的意义。

除了将色谱和质谱直接联用,一些研究小组还发展了 HPLC 和 MS 的离线联用技术。在这些技术中,HPLC 的功能大都是分离并收集样品中的特定组分,然后再借助质谱加以分析。例如,Chaurand 等[12]采用 4.6 mm × 15 cm 的常规反相液相色谱柱分别分离并收集了正常细胞和肿瘤细胞中的蛋白质,胰蛋白酶酶解以后,经 MALDI-TOF-MS 分析,成功鉴定了 3 种肿瘤专一性蛋白质。Coutre 等[13]采用 4.6 mm × 30 cm 的体积排阻色谱柱预先纯化已经经过处理的 4 种膜蛋白,以苯乙烯基质的反相色谱柱加以分离,收集到的组分经过 CNBr 处理后,通过 MALDI-TOF-MS 检测,直接测定了这 4 种不同膜蛋白的相对分子质量,得出最高相对分子质量是 61 000,这一结果的准确率达 ±0.01 %。Cunsolo 等[14]采用 1.0 cm×25 cm 的反相 C18 柱分离了酶解甲基化 *Parietaria judaica* 所得的多肽混合物,分离组分经过收集处理后进入 ESI-MS 检测,鉴定出这种花粉中引起过敏的蛋白质中二硫键的位置,同时还表征了其所含半胱氨酸的氧化态,为蛋白质和多肽中二硫键的研究提供了一个新的方法。

在一维色谱-质谱联用的技术领域中,另一个值得重视的方面是,一些研究小组报道了一些高通量筛选技术。如出现了 8 通道[15]和 96 通道[16]的液质联用系统,反相高效液相色谱和傅里叶变换离子回旋共振质谱(FTICR)的联用系统[17]等,这些也都大大促进了一维色谱-质谱联用技术在蛋白质组学中的应用。

综上所述,从一维色谱-质谱联用技术的发展可以看出,对于不是特别复杂的体系,该技术可以比较充分地发挥快速、灵敏、易于自动化的优势,在蛋白质组学的研究中占有一定的地位。

二、多维色谱及其与质谱联用技术

（一）理论基础

与一维分离模式相比，多维分离技术的最大特点是可以极大地提高峰容量。1995年，Giddings等[18]指出峰容量的提高程度和样品组分分布的有序程度有着密切的关系，并引入了样品维数(sample dimensionality)——能够精确识别样品中所有组分的变量数的概念。他们认为，当样品维数大于系统维数时，组分的峰分布是无序的，会限制分离的效果；而当样品维数小于系统维数时，尽管组分的峰分布是有序的，但多维系统高峰容量的性能并没有得到充分的发挥；只有当样品维数和系统维数相等的时候，多维系统的分离能力才能得到最大程度的发挥。同年，Giddings[19]构建了二维分离模式下实际峰容量的数学计算模型，该模型不仅能够拓展应用到评价多维分离模式的性能，也可以应用于系统的条件优化。另一方面，Murphy等[20]从理论和实验上对二维分离模式下第二维的采样速度对分离的影响做了深入研究。结果表明，进入第二维分离的采样时间越短，整个系统的选择性越高。以上这些研究为多维色谱方法的发展奠定了理论基础。以下将介绍几种不同液相色谱模式之间的组合及其应用。

（二）离子交换色谱-反相液相色谱

离子交换色谱是通过溶质在离子交换色谱固定相上具有不同的保留能力而实现样品分离的色谱技术。具体地说，生物大分子，尤其是蛋白质和多肽，表面多是带有一定电荷的，它们和离子交换色谱固定相表面的电荷可以发生不同的静电相互作用，利用这种相互作用的差异，可以实现带电生物样品的分离。而反相液相色谱是基于溶质疏水性的差异而实现分离的。通过这两种色谱模式的联用，可以实现对复杂生物样品的二维分离，这种组合也是迄今为止在蛋白质组学研究中应用得最广泛的一种联用技术之一。值得注意的是，由于离子交换色谱的流动相中含有较高浓度的盐，而反相液相色谱的流动相一般是有机溶剂和水，二者在耦联的时候就必须考虑到流动相种类不同所造成的影响。一般来说，解决这一问题常要使用到设计比较巧妙的管路和连接阀，一些小组借助这些装置，取得了一定的研究成果。例如，Holland等[21]将阴离子交换柱和RPLC微柱通过电控阀连接起来，实现了阴离子交换柱和反相柱的在线耦联，但是其阴离子交换色谱所用的分离时间长达32 h，效率比较低。Wagner等[22]将长度分别为3.5 cm和14 cm，内径均为4.6 mm的一根强阳离子交换柱和两根平行的C18反相液相色谱柱通过一个十通阀连接起来，使从IEC流出的组分可以交替进入到两根平行的RPLC柱中，实现了IEC和RPLC的在线耦联，然后通过组分收集，MALDI-TOF-MS鉴定，成功分离出了大约1 000种蛋白质和多肽。Optitek等[23]则通过一个八通阀连接起内径分别为750 μm和500 μm的阳离子交换柱和反相色谱柱，在线紫外检测，同时连接到电喷雾接口进入质谱加以鉴定，在两小时以内，成功地对大肠埃希氏菌(*Escherichia coli*)细胞中的蛋白质进行了分析鉴定。Davis[24]等通过六通阀将常规尺寸的强阳离子交换柱和RPLC耦联，分流器分流后进入串连质谱检测，如图20.2所示，对人肺纤维原细胞和人脑神经胶质瘤细胞中的蛋白质酶解产物做了鉴定，比传统方法多鉴定出了40%的肽段。但是，由于使用了分流器，一些微量组分可

能会因为损失过多而难以被检测。此外，Vissers 等[25]则通过将强阳离子交换柱分离出的组分加以收集，再进样到 RPLC 柱的方式联用了两种不同的分离模式，有效避免了一维和二维分离中流动相的混合，简化了系统的设计。

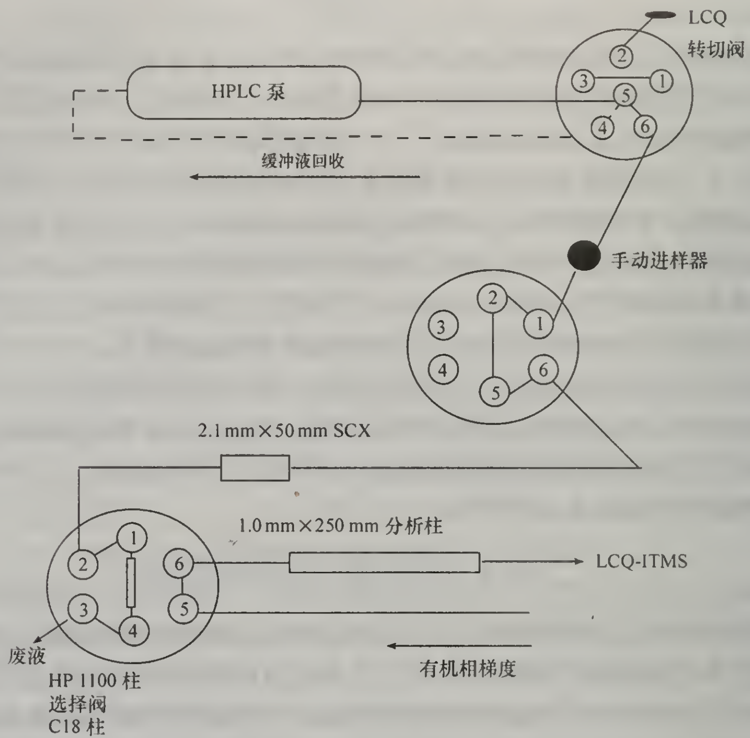


图 20.2　2D SCX HPLC-MS 联用流程图

（三）体积排阻色谱-反相液相色谱

体积排阻色谱(size exclusion chromatography, SEC)是根据体积大小不同的溶质分子，通过具有体积筛分作用的固定相时的保留程度不同而达到分离的一种色谱模式。一般说来，体积小的分子在经过体积排阻固定相时更容易扩散到固定相表面的微孔中去，因此保留较强，而体积比较大的溶质则正好相反，保留较弱。Opiteck 等[26]将 7.8 mm×300 mm 的 SEC 柱和两根平行的 4.6 mm×33 mm 的 C18 反相液相色谱柱用一个四通阀连接起来，紫外检测后通过 10∶1 的分流器引入电喷雾质谱，对卵清蛋白和血浆白蛋白的酶解产物做了分离和鉴定，如图 20.3 所示。之后，他们[27]将反相色谱柱换为 2.1 mm×33 mm 的聚苯乙烯-二乙烯基苯反相柱，采用收集最终分离组分的方法，把较好分离的大肠埃希氏菌细胞所过量表达的蛋白质送进 MALDI-TOF-MS 检测，成功进行了蛋白质的 mapping。体积排阻色谱-反相液相色谱联用模式可以提供比较高的分辨率，但是由于体积排阻色谱的成本比较高，在一定程度制约了这种技术的应用。

（四）反相液相色谱-反相液相色谱

在各种色谱联用技术中，还有一类是将两根反相色谱柱连接在一起使用的，如 Chen

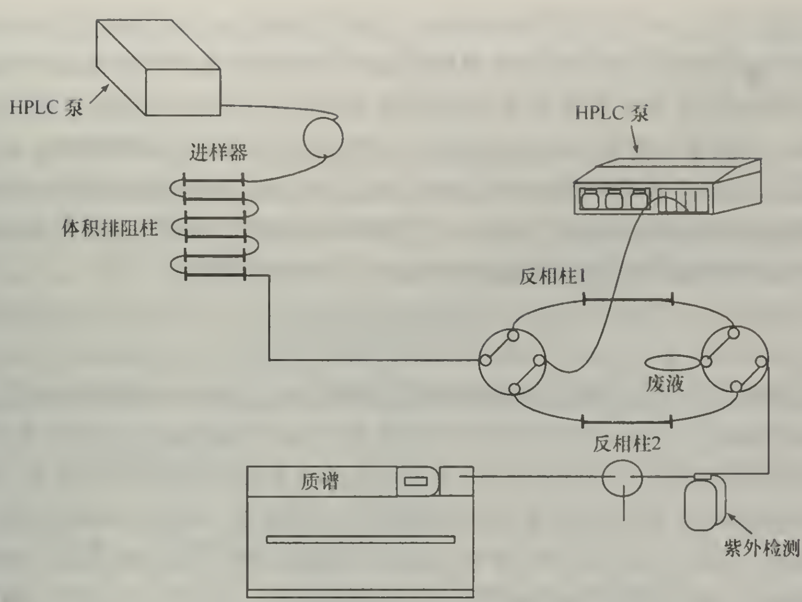


图 20.3　2D-SEC-RPLC-MS 联用流程图

等[28]使用两根尺寸不同的填充了无孔反相填料的液相色谱柱，对人红白血病细胞裂解液中的蛋白质做了分离，胰蛋白酶酶解后经 MADLDI-TOF-MS 检测，得出了这种细胞的蛋白质质量图谱。在这个系统中，第一根色谱柱预分离出了约 100 种蛋白质，然后快速通过第二根反相色谱柱，将第一相中没有分开或没有完全分开的组分继续加以分离，得到了理想的结果。而 Wall 等[29]则通过采用不同尺寸的色谱柱、填以不同粒径大小的填料，在不同的温度下连接，组成 RP-RP 连接体系，然后采用手动点样 MALDI-TOF-MS 分析了大肠杆菌表达的蛋白质，研究了大肠杆菌对树胶醛糖的反应，整个色谱分离的时间只用到了 13.5 min。值得注意的是，RP-RP 联用虽然和 Giddings 提出的多维分离技术所要求的各相分离最好是正交关系的原理有所抵触，但是从以上的应用我们可以看出，RP-RP 联用还是给复杂体系的分离提供了比较高的峰容量和选择性。

（五）亲和液相色谱-反相液相色谱

用亲和色谱柱对某一类具有特异性亲和力的蛋白质或者多肽进行提取，然后进行 MS 或 MS/MS 鉴定，确定出蛋白质的种类和修饰位点，这种方法在蛋白质组研究中也发挥了重要的作用。Apffel 等[30]利用外源凝集素柱与反相柱的在线连接以及电喷雾质谱检测对起源于重组 DNA 的糖蛋白水解后产生的不同类型的糖基化肽进行了研究，发现其中 4 个糖基化位点全部是 N 连接的。Hsieh 等[31]通过集成化的免疫亲和-离子交换-柱上酶解-片段捕获和脱盐-RPLC-MS 检测实现了蛋白质初级结构的自动化分析。此外，以外源凝集素为基质的亲和色谱和反相液相色谱以及质谱联用技术已经应用于肽的指纹谱分析中。Han 等[32]利用离子交换色谱柱将细胞微粒体中含有的蛋白质分离，收集出 30 个组分，再将这些组分进行同位素编码亲和标记（ICAT），水解后再通过自行填装的聚抗生物素蛋白的亲和色谱-反相液相色谱-电喷雾串联质谱（AC-RPLC-ESI-MS/MS）进行

分析，成功确定了 HL-60 细胞微粒体中 491 个蛋白质分子，并可以对它们在分化过程中的变化进行定量研究。Goshe 等[33]也采用类似的方法，对低丰度的磷蛋白进行富集和识别。Geng[34]则通过亲和色谱和反相色谱联用，MALDI-TOF-MS 检测，对蛋白质进行了肽指纹分析。金属螯合色谱也是亲和色谱的一种形式。近年来，随着磷酸化蛋白质组研究的不断深入，固定化金属亲和色谱也被成功地应用于蛋白质组学的研究中来，较好地解决了磷酸化蛋白质的分析和鉴定，有关内容见本书相关章节。同样，在蛋白质和多肽的糖基化研究中，亲和色谱-反相液相色谱质谱技术也具有相当潜力。

（六）三维色谱联用技术

二维色谱联用为蛋白质组学的研究提供了强有力的支持，一些研究小组为了更充分地发挥液相色谱技术在蛋白质组学研究中的优势，已经开始尝试三维液相色谱的联用。实际上，前面所提到的 Han 等人的工作，从某种意义上来说，已经可以认为是三维色谱的实际应用，他们实现了离子交换-亲和-反相液相色谱的联用，尽管第一维和后两维色谱并没有在线相连。实际上，在此之前，早在 1995 年，Moore 等[35]就已经尝试了体积排阻色谱-反相液相色谱-毛细管区带电泳联用的三维分离技术。他们认为，由于二维联用技术已经在很大程度提高了峰容量，所以第三维即使仅仅得到了一个很弱的峰，也会极大地提高整个方法的峰容量。最近，MacCoss 等[36]则成功地实现了反相液相色谱-离子交换色谱-反相液相色谱-串联质谱的三维联用技术，对蛋白质混合物同时用三种不同酶切后的肽段混合物进行了分析，成功定位出了化学计量只占 10％的磷酸化肽段的磷酸化位点，确认了总共 5 种蛋白质中的 3 种蛋白质的 4 类 18 个修饰位点，其中 3 个修饰位点是从未报道过的；同时，把这个系统应用于眼球组织蛋白质的分析中，成功鉴定出了 270 种蛋白质。其中，11 种晶状体蛋白总计被发现了 73 处修饰位点，这种方法随即被认为是鉴定蛋白质修饰位点的强大工具。为了充分展示三维色谱连用的优势，McDonald[37]则比较了一维、二维和三维液相色谱和质谱联用技术用于蛋白质混合物的分析，结果发现三维色谱的分辨率最高，鉴定出的蛋白质种类也最多。实际上，MacCoss 和 McDonald 两个研究小组所开展的工作属于另外一类多维分析技术——多维蛋白质鉴定技术（MudPIT）。

（七）多维蛋白质鉴定技术（MudPIT）

多维蛋白质鉴定技术（MudPIT），最早是 Yates 研究组[38]于 1999 年提出的。具体地说，这种技术是将不同分离模式的色谱柱以串联分离的方式合并于同一根色谱柱中进行。例如，Yates 等[38]在同一根色谱柱的前半部分装填强阳离子色谱填料，后部分装填反相液相色谱填料。这样整个色谱分离过程就包括了一系列依次增加的盐浓度台阶梯度洗脱和有机溶剂的线性梯度洗脱，多肽流出 RPLC 填料后直接进入质谱分析，最后用 SEQUEST 算法从基因组翻译的蛋白质数据库中进行检索，查找并确定与多肽序列相匹配的蛋白质。这种方法已经被成功地用于蛋白质组学中的研究，并被称为多维蛋白质鉴定技术（multidimensional protein identification technology，MudPIT）。采用这种方法，有效避免了复杂的阀切换系统，减小了系统的死体积，而且由于该技术是将强阳离子和反相液相色谱填料依次装填在一个纳喷进样针中，一个分析周期可以检测 100 多种蛋白质，所以该方法可对样品量较少的蛋白质进行快速分析，适用于蛋白质组学中大规模蛋白质的分离鉴定。将

该技术用于 Saccharomyces cerevisiae strain BJ5460 grown to mid-log phase 的蛋白质组研究中，共检测和识别出 1 484 种蛋白质，其中包括低丰度蛋白质（如转录因子和蛋白激酶）和 131 个跨膜区域的蛋白质[39]。分离模式如图 20.4 所示。之后，Wolters 等人[40]改进了这项技术，所有识别的蛋白质所跨越的动态范围可以达到 10 000∶1。如前文所述，他们利用三维固定相柱系统实现了 3D-RPLC-IEC-RPLC 对在转录和后转录阶段蛋白质修饰的研究。在对蛋白质组进行定性分析的同时，以 MudPIT 技术对其进行的定量研究也在开展。Washburn 等[41]将代谢标记引入 MudPIT，他们首先对酿酒酵母菌株在富含^{14}N或^{15}N的培养基中分别培养，然后将两者的裂解液按不同的比例混合并酶解成多肽混合物，用 MudPIT 分析混合物中的蛋白质组，获得了准确、可靠的结果，表明该方法可用于一系列微生物，甚至哺乳动物组织培养体系蛋白质组的定量研究。

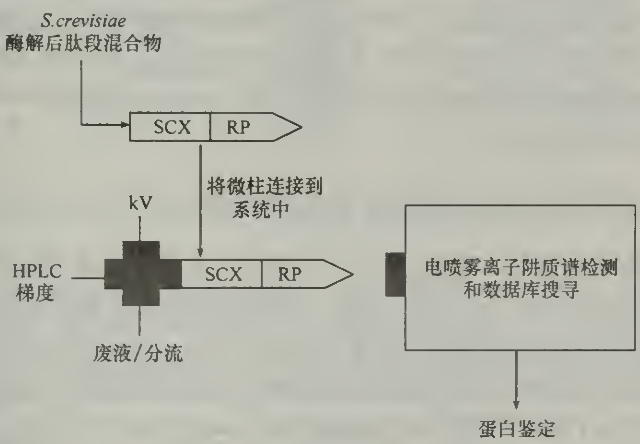


图 20.4　MudPIT 用于酿酒酵母（*Saccharomyces cerevisiae*）BJ5460 株蛋白质组研究中的技术路线

第三节　非凝胶电泳

一、液相等电聚焦

（一）简　　介

等电聚焦（isoelectric focusing，IEF）最早由 Svensson[42~45]等提出，在聚焦小室中，蛋白质样品与特定 pH 范围内的载体两性电解质混合物（carrier ampholyte），或其他缓冲液混合，当向聚焦小室施加一个电势能时，蛋白质就会迁移到已经建立好的 pH 梯度中的一个特定的位置，这个位置的 pH 与它们各自的等电点（p*I*）相当。如果一个蛋白质的 pH 不在这个范围内，它的净电荷将发生改变，随之产生的电泳驱动力将影响其迁移，使其回到其等电点所在的位置上。最理想的结果是蛋白质聚焦到它们的等电点位置并形成狭带。液相等电聚焦有许多优点：①能够大量纯化蛋白质和多肽，尤其是在相当高的上样量时也具有相当高的分辨率；②可在非凝胶介质上根据等电点将蛋白质的复杂混合物分级，每一级都能被收集，并且通过电泳或色谱进一步分析；③应用等电聚焦分离生物样品时（如蛋

白质混合物),生物样品的活性易于保持[46];④对于低丰度蛋白质和低相对分子质量的蛋白质也具有很好的分离效果[46,47];⑤可以避免在双向凝胶电泳分离状态下存在的重复性差的问题[48]。IEF 的不利之处在于聚焦后高浓度的中性蛋白质常会从溶液中沉淀出来,造成组分间的相互重叠;用于构建 pH 梯度的两性电解质可能会干扰后续的某些分析(如 ESI-MS);高疏水性蛋白质在样品制备或蛋白质到达等电点后继续聚焦的过程中可能会丢失。

能够将复杂混合物分级到备用水平的 IEF 仪器有多种:①Rotofor cell 仪器(Rotofor Cell Apparatus)[48];②多腔等电膜电解器(multicompartment electrolyzer with isoelectric membranes)[49];③Bier[50~53]首创的循环等电聚焦仪(recycling isoelectric focusing, RIEF; RF3);④自由流等电聚焦(free-flow IEF, FFIEF)[54]。在 RIEF 中,样品介质通过冷却室循环,而 FFIEF 中样品被持续注入作为载体的两性电解质溶液中。复杂的蛋白质溶液很容易被分级为很多组分,它们再被收集,并应用正交技术进一步分级。关于应用等电聚焦分级技术对细胞和血清中的蛋白质进行分离的研究,已有相关报道[54~56],下文将具体介绍。

(二)循环等电聚焦在酵母全细胞裂解液样品分析中的应用

循环等电聚焦最早由 Bier 等提出[50~53],用于在液相中大量制备样品(如蛋白质)。现以 RF3(recycling free-flow focusing)为例,介绍利用循环等电聚焦来对蛋白质混合物进行分级分离。RF3 由 Bier 等人设计,由美 Protein Technologies 公司生产。如图 20.5 所示。在循环等电聚焦过程中,首先,载体两性电解质先要在聚焦室中聚焦,并在聚焦室内形成一系列固定的 pH 梯度(与载体两性电解质的 pH 范围相当),然后,把蛋白质混合物注射到聚焦室里,蛋白质混合物就会随着载体两性电解质溶液在 RF3 的循环管路中循环往复地流动。当蛋白质流动到聚焦室里时,由于载体两性电解质在聚焦室里形成了固定的 pH 梯度,蛋白质会逐渐向与其 p*I* 值相当的位置迁移,并最终到达那里,形成一条狭带,稳定地在这一区间循环流动,直至分离完成,并被收集下来。复杂的蛋白质混合物根据其等电点被分成了多个组分。每一组分中虽然还是以蛋白质混合物的形式存在,但是,其复杂度已经小多了。循环等电聚焦的另一优点,即分离是在液相中进行的,可以直接进行溶液内酶解,然后,使用 ESI-MS 鉴定每一组分的蛋白质混合物(shotgun, 鸟枪法)[57,58],继而,合并多个 RF3 组分中所鉴定出的蛋白质,就可以得到所要鉴定的总蛋白质混合物。此外,还可以把 RF3 分离得到的产物进行如聚丙烯酰胺凝胶电泳分离,再进行质谱分析。利用这一实验流程,我们可以鉴定细胞或组织的蛋白质组。下面将以本实验室的工作为例,说明具体研究步骤。

1. 制备酵母全细胞裂解液样品

酿酒酵母 YPH499 由中国科学院上海生命科学学院生物化学与细胞生物学研究所分子生物学重点实验室马普四组获得。酵母细胞的培养参照《分子克隆实验指南》第 3 版[59]。

(1) *酵母细胞的培养*

YPD 是一种复杂的培养基,用于常规的酵母的生长,YPD(YEPD) 培养基成分为酵

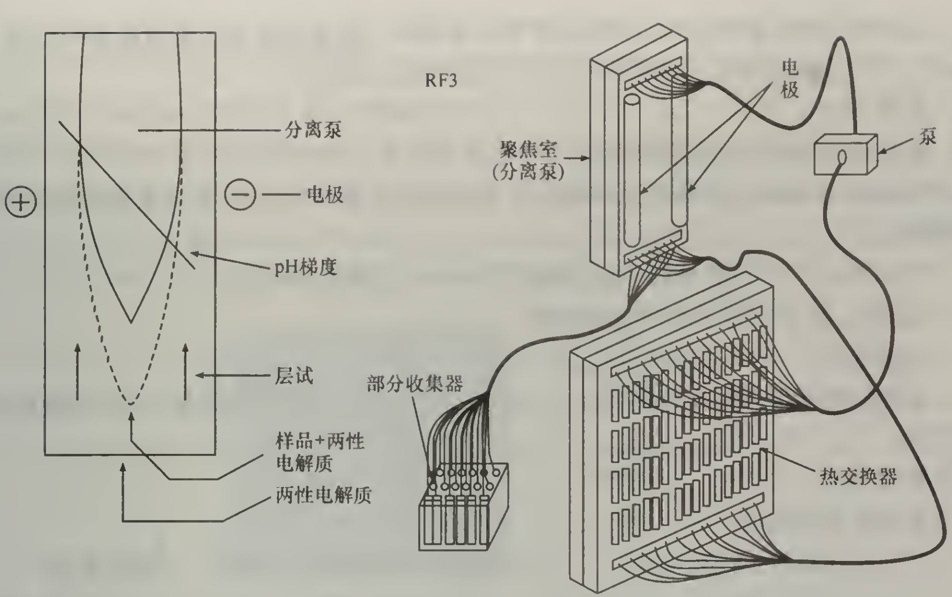


图 20.5　RF3 原理图和装置

母抽提物(yeast extract, 10g),peptone(20g),葡萄糖(Glucose, 20g),加如 Millipore 级去离子水至 1 000 ml。用灭菌过的牙签挑一个克隆,放在 50 ml YPD 培养基中于摇床上培养过夜,然后再加入 YPD 培养基至 2 L,置于摇床上摇过夜。收集在对数期生长的酵母细胞。

(2) 裂解酵母细胞

(A)把约 2 L 的酵母溶液倒入 4 个离心管中,4℃、6 000 *g* 离心 5 min。

(B)除去上清液,此时酵母细胞沉淀在离心管的底部。

(C)用去离子水洗涤沉淀,以除去细胞表面的盐分。

(D)4℃、6 000 *g* 离心 5 min,并除去上清液。

①加入约 50 ml 裂解液(1%载体两性电解质+10%甘油)至离心管中,用力摇使沉淀悬浮,然后,使用高压裂解酵母细胞(物理裂解)。

②4℃、6 000 *g* 离心 5 min,把上清液倒入 50 ml 管中(约 30 ml)。

③4℃、15 000 *g* 离心 1h。

④取上清液 4℃、40 000 *g* 离心 1 h。

(3) 蛋白质定量分装及保存

采用 Bradford 法定量酵母裂解液中的蛋白质浓度,并把酵母裂解液分装到 1.5 ml 的 Eppendorf 管中,每管 1 ml。把分装到 Eppendorf 管中的酵母裂解液置于液氮中,数分钟后,再放在 −80℃ 低温冰箱中保存备用。

2. 用 RF3 进行分级分离

(1) RF3 相关的电泳液组分

①电极液:“+”:0.1 mol/L H_3PO_4;“−”:0.1 mol/L NaOH。

②介质溶液：使用 RF3 进行循环等电聚焦时，等电聚焦在 1% 的载体两性电解质（pH3.5～10）和 10% 的甘油溶液中进行。总体积为 105 ml。

(2) 清洗 RF3

①500～1 000 ml 去离子水。

②300～400 ml 0.1 mol/L NaOH，并在管道那保留约半小时（主要是清洗收集器和聚焦管路）。

③1 000～1 500 ml 去离子水清洗除去 NaOH 及蛋白质残液。

④排空，并以去离子水冲洗电极槽。

⑤放松泵。

⑥短期内不使用时在管路内注入 10% 的乙醇，若长期不使用则加入 5% 的叠氮钠，防腐之用。

(3) 分离过程

①注入介质溶液。

（Ⅰ）打开冷凝系统预冷至 2℃，在电极槽中分别加入“+”极和“-”极电极液。

（Ⅱ）打开排气管的阀门，开主泵，“pump direction”处于“fill”位置，泵速为：20%～30%。

（Ⅲ）注入介质溶液直到排气管中的溶液至 2～3 ml（～70 ml）。

（Ⅳ）停止“pump”。

（Ⅴ）连续注入介质溶液至聚焦室顶部，停止注入。

（Ⅵ）打开“pump”，直到聚焦室里的介质溶液全部流入管路里。

（Ⅶ）停止“pump”，重复（Ⅴ）和（Ⅵ）步骤的操作，直至 105 ml 介质溶液全部被注入到循环管路里。

（Ⅷ）打开“pump”，调节“pump”的位置至“Run/collect”；泵速在 50%～60%，敲击聚焦室以排出气泡。

②预聚焦过程。

设定“pump”的位置至“Run/collect”，泵速在 45%。

1 500 V，400 mA，150 W，使载体两性电解质先在聚焦室内进行聚焦，形成一系列固定的 pH 梯度，并在封闭的管道内循环流动。

③当电流降低并稳定后，加入蛋白质溶液（酵母裂解液）20 mg，并降低高压至 1 000 V。

④聚焦过程（1.5 h，蛋白质/载体两性电解质溶液）。

（Ⅰ）1 000V，1h（在电流降低并稳定后）。

（Ⅱ）降低电压至 500 V，并持续半小时。

⑤同时关掉“pump”和高压；然后关上排气阀。

⑥快速转移“inlet connector”至收集器，并连接“voirse tube”至聚焦室底部。

(4) 收集 RF3 产物

①打开排气阀，使排气管和聚焦室里的介质溶液和蛋白质混合物由于重力的作用经“voirse tube”流出。

②打开“main pump”，收集分离产物至 30 个收集管里。

3.RF3 产物结果的分析及处理

(1) 测定各个组分的 pH 值及蛋白质浓度

RF3 产物被收集到收集管里后，使用微量 pH 计测量各个管里组分的 pH，30 管 RF3 产物的 pH 从高到低排列。如图 20.6 所示。各个组分的蛋白质浓度可以用 OD_{280}来进行粗测，其中紫外吸收值(ABS)1.0 约相当于 1.0mg/ml。另外用 Bradford 法可以更精确地测定各个组分的蛋白质浓度。如图 20. 6 所示。

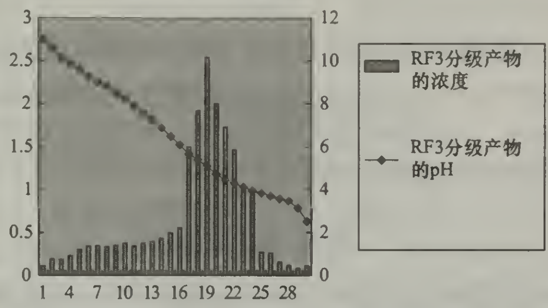


图 20.6　酵母裂解液的 RF3 分级产物的 pH 和浓度示意图

(2) 重复性分析

循环等电聚焦的优点之一，即可以避免在双向凝胶电泳分离状态下存在的重复性差的问题[48]。表 20.1 所示即为 3 次使用 RF3 来分级酵母裂解液的 pH 列表。可见，被收集在同一收集管中的 RF3 分级产物其 3 次测定的 pH 相差＜0.1，大多数都在 0.05 以下。这说明循环等电聚焦的重复性是相当好的。

表 20.1　RF3 分级产物的 pH 的重复性分析，表中所示为 3 次使用 RF3 来分离酵母裂解液的 pH

管号	pH			管号	pH			管号	pH		
1	11.03	11.01	11.01	11	8.06	7.87	7.92	21	4.43	4.46	4.47
2	10.67	10.59	10.59	12	7.59	7.56	7.56	22	4.25	4.29	4.29
3	10.30	10.32	10.15	13	7.27	7.26	7.25	23	4.11	4.10	4.12
4	9.99	10.02	9.87	14	6.89	6.77	6.88	24	3.95	3.93	3.96
5	9.75	9.70	9.63	15	6.45	6.31	6.49	25	3.85	3.85	3.85
6	9.41	9.40	9.26	16	6.02	5.97	6.09	26	3.70	3.68	3.70
7	9.10	9.10	9.02	17	5.60	5.51	5.67	27	3.56	3.57	3.58
8	8.95	8.85	8.82	18	5.40	5.35	5.39	28	3.44	3.52	3.48
9	8.54	8.56	8.50	19	5.05	5.09	5.10	19	3.17	3.22	3.15
10	8.25	8.18	8.24	20	4.69	4.70	4.73	30	2.65	2.59	2.52

4. 讨论

除了上述 RF3 以外，还有其他一些用于循环等电聚焦的仪器来分离蛋白质混合

物[60~65]，如 Tattanahalli[60]等所应用的一种等电聚焦仪，其自动化程度更高。在分离过程中实现了计算机控制，并且，各个组分的 pH 值和浓度也可以在线检测。使得实验过程更加精确。使用 RF3 分级后，再把蛋白质溶液在反相液相色谱中进行分离，也是一个很好的策略，经液相等电聚焦和反相色谱分离后，使用一种软件可以把蛋白质模拟成二维胶的分离模式[61,62]。

二、毛细管电泳

与等电聚焦类似，毛细管电泳也是一项具有很高分离效率的分离技术，20 世纪 90 年代，这种技术被广泛应用于 DNA 分析。人类基因组测序最终能够提前完成，很大程度上也要归功于这种分离技术。毛细管电泳的基本原理在本书的前面章节已经有所介绍，以下将主要讨论其在蛋白质组学研究当中的应用。

（一）毛细管区带电泳

毛细管区带电泳是毛细管电泳各种模式中应用最为广泛的微分离技术，对肽段和蛋白质的分离中可达到 10^6理论塔板高度。这意味着毛细管区带电泳的峰容量应该是 10^3数量级，但由于缓冲溶液的 pH 是相对固定的，不可能所有的肽段等电点都在这个 pH 的附近。所以，实际上，毛细管区带电泳一次可以分开的肽段并没有理论上这么大。但是，一些研究小组还是利用这种技术进行了相关的研究，不过绝大多数研究对象都相对简单。例如，He 等[66]将两种乳球蛋白用胰蛋白酶、胃蛋白酶和胰凝乳蛋白酶柱上酶解在 20 支毛细管中，在 40 min 内，以 MKEC 的方式，二极管阵列检测，检测出了这两种蛋白质的酶解肽段。实际上，由于毛细管电泳与质谱接口之间的技术问题还没有得到很好解决，所以毛细管电泳和质谱联用技术在蛋白质组学研究中的应用也在一定程度上受到了限制。不过，Bindila[67]等以甲酸胺为缓冲溶液，在低 pH 下分离了几种标准蛋白质的酶解产物，然后电喷雾-四极杆-飞行时间串联质谱分析，成功鉴定出了这些蛋白质。Wetterhall[68]等人采用毛细管区带电泳和电喷雾-傅里叶离子回旋共振加速器联用，对胰蛋白酶酶解脑脊液所得的肽段做了分析，结果在 95% 的可信度和小于 5 ppm 的质量测量误差范围内，鉴定出了 30 种蛋白质。有关 CE-MS 的综述见参考文献[69,70]。

（二）毛细管等电聚焦

毛细管等电聚焦已经被应用了近 30 年，也被认为是所有一维分离方法中具有最高分离效率的手段。Shen[71]对胰蛋白酶酶解酵母胞液所得的样品进行分析，在 0.045 cm（毛细管总长）的区带宽度内得到了半峰宽小于 3 s 的样品峰，最小区分等电点 ΔpI 是 0.005，峰容量接近 1 000。但是，40 min 的分离时间显得耗时稍长。接着，在对 *D. radiodurans* 细胞裂解液中的蛋白质进行分析的过程中，他们又报道了峰容量大约在 800 的 CIEF 分离[72]。结果在 pH 3.0～8.8 范围内，鉴定出了 210 个样品峰，ΔpI 达到 0.004。Manabe 等[73]在内径 75 μm、长 27 cm、内壁键合了聚丙烯酰胺的毛细管内，以 CIEF 的方式分离了人脑脊液中的蛋白质混合物，鉴定出了 70 种蛋白质。与此同时，一些研究小组还研究了毛细管管壁改性情况对分离效率的影响。Shen 等[74]发现，亲水性的内表面更适合 CIEF

对蛋白质的分离，他们先使用了内表面固定纤维素的毛细管作为等电聚焦的载体，发现最小区分等电点为 0.018～0.013，而当使用线性聚丙烯酰胺改性的毛细管在 pH 3～10 范围内，对 *Deinococcus radiodurans* 中的蛋白质进行分离时，最小区分等电点则达到了 0.004，峰容量为 800。值得指出的是，为了避免样品沉淀，CIEF 所用的样品浓度一般小于 0.1 μg/μl，为了增大上样量，一些研究小组[74]采用了相对长度较长、内径较粗的毛细管，在不显著影响分离效率的前提下，上样量也达到了 10 μl 以上，也就是 10 μg 左右，也得到了较好的分离结果。同时，由于 CIEF 是在毛细管内进行的等电聚焦，所以，对采用 2D-PAGE 无法分离的超大蛋白质也具有分离能力。因此，除了对多肽和蛋白质的分离，CIEF 还成功地应用于蛋白质复合体的分离，相关进展见参考文献[75～77]。

（三）多维毛细管电泳

与多维液相色谱相比，多维毛细管电泳在蛋白质组研究中的应用正处于起步阶段。Sheng 等[78]利用 MEKC（胶束电动色谱）和 CIEF 在线联用，在 4min 内完成了对一种蛋白质水解产物的浓缩、分离和检测。系统通过十通阀对一维的流出组分进行收集和脱盐处理，同时与和 CIEF 连接的八通阀相连作为接口，如图 20.7 所示。Mohan[79]等则通过微透析界面实现了 CIEF 和 CZE 的联用，并将该模式应用于蛋白质水解肽的分析中，峰容量可以达到 1 600 左右。由于在不同模式的毛细管电泳技术切换流动相还有不少技术性的难题未解决，因此多维毛细管电泳的使用还处于探索阶段。

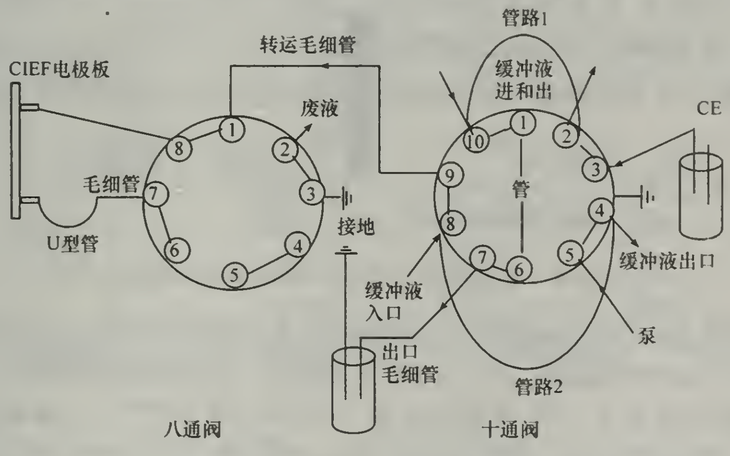


图 20.7　2D-MEKC-CIEF 分离模式路线图

（四）毛细管电泳-液相色谱联用

多维 HPLC-CE 分离系统最初是由 Jorgenson 等于 1993 年[80,81]提出的，当时他们设计了两种接口，即阀/环设计（valve/loop design），如图 20.8 所示；横向流动界面（transverse flow gating interface），如图 20.9 所示。

后来，他们又利用这两种接口实现了 2D-RPLC-CZE[82,83]和 3D-SEC-RPLC-CZE[35]。1997 年，他们改进了 HPLC 和 CE 的接口部分[84]，但是进样过程中的样品稀释和堆积样品需要流出液较多的问题比较突出。Chen 等[85]以微进样器作为接口建立了 CIEF 和毛

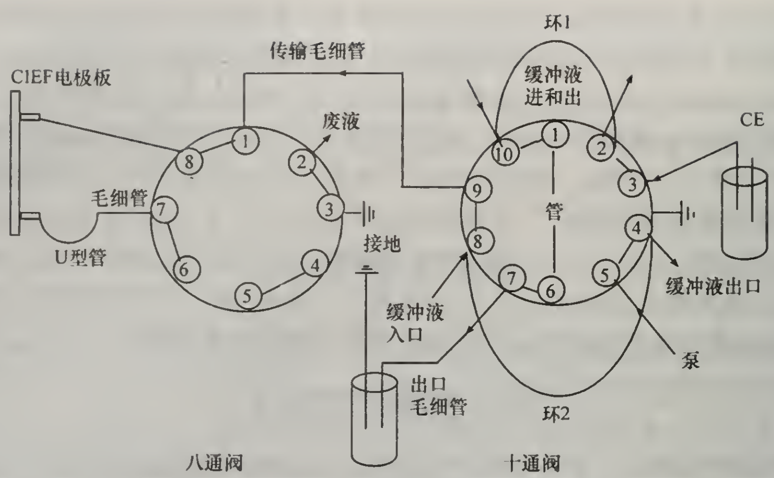


图 20.8　2D HPLC-CE 联用中的阀、环设计

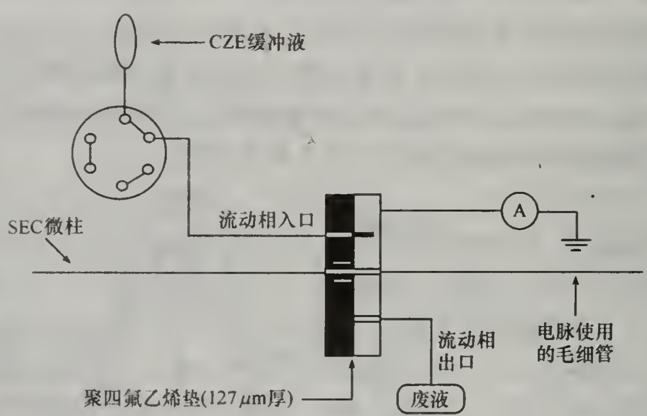


图 20.9　2D HPLC-CE 联用中的横向流控界面设计

细管反相液相色谱的二维分离系统，在对果蝇唾液腺中提取的多肽样品进行分析时，可以在 8 h 内获得 1 800 的峰容量。而且实验中还观察到了不同时期提取的样品之间的差异。在磷酸化肽段的研究中，CZE 也成功地和 IMAC-MS 联用[86,87]，检测限可以达到 28 fmol，是 2D-PAGE 的 1/20。在离线分析方面，Figeys 等[88]以 IMAC-SPE-CZE-MS/MS 实现了对蛋白质后转录阶段修饰的研究。Issaq 等[89]也实现了 HPLC 和 CZE 的离线多维分离。采用这种方法不仅可以对 HPLC 的流出组分进行浓缩，而且还可以通过 96 通道 CZE 来缩短蛋白质水解产物和细胞提取物的分析时间。Tong[90]等采用了如图 20.10 所示的设计实现了 LC-CE-MS 的联用，对酵母核糖体上的蛋白质做了分析鉴定。具体做法是：首先，样品酶解以后被酸化，得到的肽段在填充了 ODS 的 SPE 预柱内富集，然后用有机溶剂的台阶式梯度洗脱将肽段冲洗至内壁已有涂层的毛细管内，CE 模式分离后进入质谱检测。

（五）毛细管电泳芯片

微型化分析仪器由于具有高效率、高通量、试剂消耗少等优点，近年来发展十分迅速。

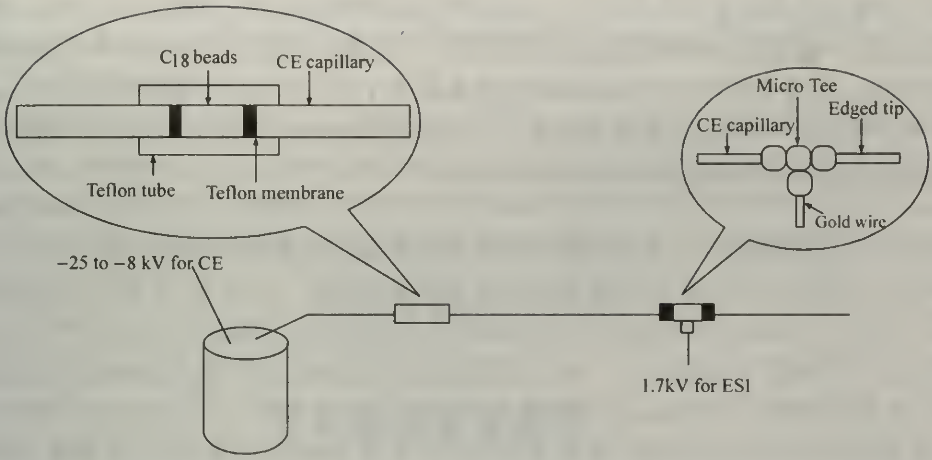


图 20.10　2D HPLC-CE-MS 用于酵母核糖体蛋白质酶解肽段分析的技术路线

毛细管电泳芯片最早是由 Hofmann 等[91]于 1999 年设计并建立的。他们在玻璃板上构造了 200 μm 宽、10 μm 深、7 cm 长的多通道，对已标签的多肽和蛋白质进行了等电聚焦分离，5min 内峰容量达到了 30～40，理论塔板高度为 0.4μm。在此以后，一些研究小组[92~94]通过特殊的死体积很小的连接装置实现了毛细管电泳芯片和电喷雾质谱的联用，示意图见图 20.11Gottschlich[95]等将蛋白质的在线酶解、分离、柱后衍生等技术引入了毛细管电泳芯片，在总计 15 min 的分析时间内，以 1 min 为分离时间，完成了对胰蛋白酶酶解的胰岛素的氧化 B-Chain 的分析。Ramsey[96,97]的研究小组在玻璃芯片上将

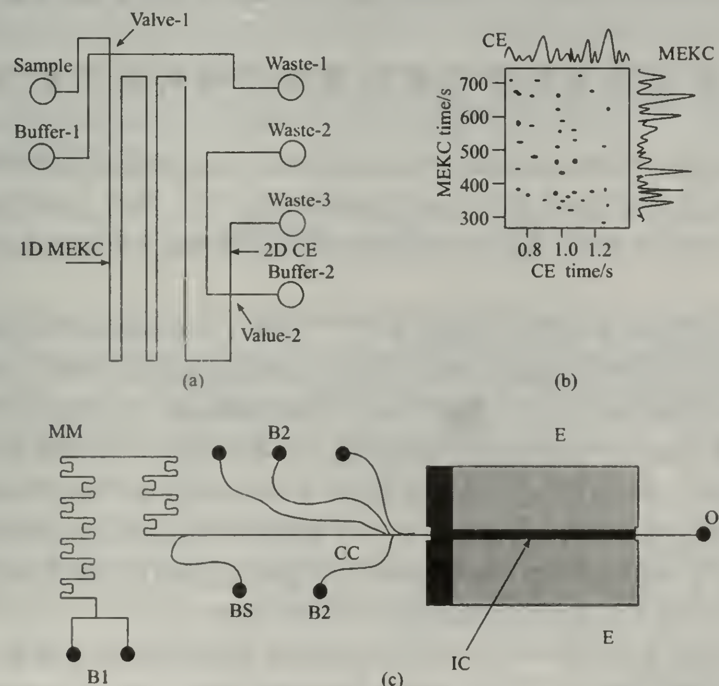


图 20.11　芯片上的多维毛细管电泳分离技术

(a)2D MEKC-CZE；(b)2D MEKC-CZE 对 Bovine hemoglobin 酶解肽段的分析；(c)2D IEC-CIEF

MEKC 的流出组分快速引入 CZE 中,实现了时间上的多维分离,如图 20.11 所示。最近他们通过进一步优化分离通道的拐角形状和尺寸、提高一维组分的采集速率及减少进样区带的方法,将整个系统对肽分离的峰容量提高到 4 500,分析时间液缩短至 15 min 内。2003 年, Wang[98]等通过在毛细管电泳芯片上使用 interferometric backscatter 检测,将未标签的蛋白质的检测灵敏度提高到 attomole 级。

第四节　毛细管电色谱及其在蛋白质组学研究中的初步应用

一、毛细管电色谱简介

毛细管电色谱是一项新兴的微分离技术,最初是通过在毛细管内填充液相色谱填料,在毛细管两端加以高电压,以电渗流驱动流动相,根据溶质在固定相上的保留强弱和电泳淌度的大小不同而实现分离。由于毛细管内有色谱固定相,电渗流驱动的流动相流型是楔子型的,因此,这种微分离技术在具有高效液相色谱的高选择性的同时,又具有毛细管电泳的高效性,已成为分析化学界近年来比较关注的领域。根据毛细管内固定相种类的不同,毛细管电色谱又有开管、填充和原位柱之分。毛细管电色谱诞生之初的应用领域主要是对中性小分子化合物的分离,但是随着这种技术在方法学上的逐步完善,更多的分离模式和固定相被相继采用。当前,毛细管电色谱在生物学领域内也有了初步的应用,但只局限于比较简单的样品,在此简介这种微分离技术在蛋白质和多肽分离方面的应用。

二、毛细管电色谱在蛋白质和多肽分离中的应用

蛋白质和多肽等生物样品,表面带电情况比较复杂,而毛细管电色谱的固定相表面又是带有一定的正电荷或者负电荷的(否则电渗流将无法产生),两者之间可能会存在一定的吸附作用。因此,尽量减少甚至避免这种静电相互作用就成为毛细管电色谱分离生物样品的关键。

电色谱中最常用的是填充柱,Unger 等[99~101]研究了一系列含苯环的肽段在以硅胶为基质的烷基反相毛细管电色谱柱和在固定相上含有磺酸基团的毛细管反相柱上的保留行为。为了减小吸附,整个分离过程使用 pH 较低的缓冲溶液,并在流动相中添加了较高浓度的盐,结果引起了不可忽视的焦耳热效应,且分离时间偏长。Ye 等[102]以强阳离子交换色谱固定相作为毛细管电色谱的固定相,在 3.5 min 内分离了 10 种小肽,最高柱效达到 46 万,但是这些小肽的长度大多为 2～3 个氨基酸,对蛋白质组学研究暂时还没有太大意义,而且由于和强阳离子表面静电吸附作用的缘故,这种模式对碱性小肽没有分离能力。后来,Ye 等[103]又采用动态改型电色谱成功分离了碱性小肽。

在开管柱方面,Pesek 等[104]以十八烷基和胆固醇基修饰毛细管内表面,分离了一些简单的小肽和蛋白质混合物。Huang 等[105]将氯甲基苯乙烯－二乙烯基苯聚合物引入毛细管内表面,然后再对聚合物改性,引入季胺和 C12 基团,以 PLOT(porous layer open

tublar)CEC 模式分离了肽段的混合物,结果显示比 CE 有更高的分离效率。Xu 等[106]则在毛细管内表面引入固定化的天冬氨酸,等度洗脱了几种标准蛋白质的混合物,超过了 HPLC 梯度洗脱的效果。

在原位柱方面,1997 年,Palm 和 Novotny[107]在 C12 衍生后的多孔聚丙烯酰胺-乙烯基乙二醇-丙烯酸原位柱上等度洗脱分离了 5 种含 Tyr 的长度不同的肽段;而实际上,为了避免样品在固定相上的吸附,使用原位柱对肽段和蛋白质的分离主要是以"Counter-direction"模式进行的,即采用表面带正电荷的固定相,在低 pH 下进行分离。这样电渗流的方向是流向正极的,而低 pH 下肽段或者蛋白质的电泳方向是流向负极的。Ericson 和 Hjerten[108]在 pH 2.0 条件下在表面带有季胺盐的和 C18 基团的原位柱上分离了几种标准蛋白质。Horvath 等[109]则通过衍生化使聚苯乙烯 - 二乙烯基苯表面带有三级胺和 C18 基团,然后在 pH 3.0 条件下分离了肽段,之后再衍生三级胺成为季胺盐,分离了核糖核酸酶、胰岛素、乳白蛋白、肌血球素和 4 种血管紧缩素。在此之后,他们在氯甲基苯乙烯 - 甲级丙烯酸乙二醇酯的整体柱上引入丁基季胺盐,等度洗脱条件下成功得出了胰蛋白酶酶解的细胞色素 C 的肽谱[110]。除了这种模式以外,Wu 等[111,112]还报道了以反相-离子交换混合模式和以电泳驱动的电色谱模式分离小肽,从方法学上丰富了 CEC 的模式。

在 CEC 芯片方面,Regnier 等[113,114]通过阵列的"配置原位支持结构"(collocate monolith support strcture,COMOSS),以 CEC 模式分离了胰蛋白酶酶解的标准蛋白质肽段。接着,Slentz 等[115]又将 COMOSS 引入了聚二甲氧硅烷(PDMS),在这种芯片上,荧光标记的 BSA 胰蛋白酶酶解产物被成功分离,柱效达到 40 万/m。2003 年,Slentz 等[116]将金属螯合色谱固定相和反相固定相引入 COMOSS 芯片,并巧妙地解决了塞子(frits)的制作问题,成功实现了标准蛋白质经胰蛋白酶解后含组氨酸残基肽段的分离。

另外,近年来 CEC 和质谱联用以及梯度洗脱方面也取得了进展。从以上对电色谱在多肽和蛋白质分析领域内应用的情况可以看出,这种技术还有待于进一步发展,才能适应蛋白质组学研究的需要。

第五节　展　　望

蛋白质组学的迅猛发展给分离科学特别是分析化学提出了日益紧迫的发展要求,新技术、新方法、新手段也就应运而生。原有的双向凝胶电泳技术虽然在蛋白质的分离纯化方面仍具有不可比拟的优势,但是非凝胶分离技术,现在已经不能被简单地认为是双向凝胶电泳这种古老的分离方法的补充了。在这些非凝胶分离技术当中,多维分离模式联用地诞生,为复杂样品的分离提供了更好的选择性和更高的分离效率。但是由于这些模式之间本身存在着一定程度的不兼容或难兼容,在一定程度上制约这些方法的组合和使用。因此,各种分离模式之间的接口选择和设计就成了这些技术进一步发展的关键。

随着材料学的不断发展,一些功能更强的新型分离介质和分离方法也必将在蛋白质组学的研究领域内发挥重要作用。原位柱由于其内部孔径结构易于控制,表面化学性质也易于修饰,所以可以按照分离目的的不同合成具有特殊孔结构和表面化学性质的有机或无机聚合物,具有广泛的应用前景。而且,就地聚合的特点也为其在芯片方面的应用铺

平了道路。

芯片技术具有高通量、高效率、成本低等优势，一旦在蛋白质组学方面获得大规模应用，对整个生命科学，将起到意想不到的推动作用。但是目前的蛋白质芯片技术实际上还处于初步发展时期。虽然化学腐蚀、光刻技术等微机电加工技术为芯片的制造提供了便利条件，但是这些技术一般实验室还难以掌握，也在一定程度上制约了其发展。

总之，非凝胶分离技术在蛋白质组学研究中正在发挥着越来越重要的作用，相信在不久的将来，随着各种技术的不断发展和完善，这些非凝胶分离技术将会更大程度地促进蛋白质组学的研究。

（靳文海　李荣霞　邹汉法）

参考文献

1 Davis MT, Stahl DC, Hefta SA *et al*. A microscale electrospary interface for on-line, capillary liquid chromatography/tandem mass spectrometry of complex peptide mixtures. Anal Chem, 1995, 67: 4549～4556

2 Moore RE, Licklider L, Schumann D *et al*. A microscale electrospray interface incorporating a monolithic, poly (Styrene-dininylbenzene) support for on-line liquid chromatography/tandem mass spectrometry analysis of peptides and proteins. Ana Chem, 1998, 70: 4879～4884

3 Premstaller A, Oberacher H, Walcher W *et al*. High-performance liquid chromatography-electrospray ionization mass spectrometry using monolithic capillary columns for proteomic studies. Ana Chem, 2001, 73: 2390～2396

4 Skodld K, Svensson M, Kaplan A *et al*. A neuroproteomic approach to targeting neuropeptides in the brain. Proteomics, 2002, 2: 447～454

5 Blonder J, Goshe MB, Pasa-Tolic L *et al*. Enrichment of integral membrane proteins for proteomic analysis using liquid chromatography-tandem mass spectrometry. J Prot Res, 2002, 1: 251～360

6 Spahr CS, Davis MT, McGinley MD *et al*. Towards defining the urinary proteome using liquid chromatography-tandem mass spectrometry I. Profiling an unfractionated trypitic digest. Proteomics, 2001, 1: 93～107

7 McCormack AL, Schieltz DM, Goode B *et al*. Direct analysis and identification of priteins in mixtures by LC/MS/MS and database searching at the low-femtomole level. Anal Chem, 1997, 69: 767～776

8 Zhou H, Ranish JA, Watts Jd *et al*. Quantitative proteome analysis by solid-phase isotope tagging and mass spectrometry. Nat Biotechnol, 2002, 19: 512～515

9 Vissers JP, Blackburn RK, Moseley, MA. A novel interface for variable flow nanoscale LC/MS/MS for improved proteome coverage. J Am Soc Mass Spectrom, 2002, 13(7): 760～771

10 Shen Y, Zhao R, Berger SJ *et al*. High-efficiency nanoscale liquid chromatography coupled on-line with mass spectrometry using nanoelectrospray ionization for proteomics. Anal Chem, 2002, 74: 4235～4249

11 Laurell T, Marko-Varga G, Ekstrom S *et al*. J Biotechnol, 2001, 82(2): 161

12 Chaurand P, DaGue BB, Pearsall RS *et al*. Profiling proteins from azoxymethane-induced colon tumors at the molecular level by matrix-assisited laser desorption. ionization mass spectrometry. Proteomics, 2001, 1: 1320～1326

13 Coutre J, Whitelegge JP, Gross A *et al*. Proteomics on full-length membrane proteins using mass spectrometry. Biochem, 2000, 39: 4237～4242

14 Cunsolo V, Foti S, Saletti R *et al*. Detection and localization of disulphide bonds in a synthetic peptide reproducing the sequence 1-30 of Par j 1.0101 by electrospary ionization mass spectrometry. Proteomics, 2001, 1: 1043～1048

15 Feng B, Patel AH, Keller PM *et al*. Fast characterization of intact proteins using a high-throughput eight-channel parallel liquid chromatography/mass spectrometry system. Rapid Comm in Mass Spectrom, 2001, 15: 821～826

16 Watt AP, Morrison D, Locker KL *et al*. Higher throughput bioanalysis automation of a protein precipitation assay using a 96-well format with detection by LC-MS/MS. Anal Chem, 2000, 72: 979～984

17 Shen Y, Zhao R, Belov ME *et al*. Paked capillary reversed-phase liquid chromatography with high-performance electrospray ionization fourier tranform ion cyclotron resonance mass spectrometry for proteomics. Anal Chem, 2001, 73: 1766～1775

18 Giddings JC. Sample dimensionality: a predictor of order-disorder in component peak distribution in multidimensional separation. J Chromatogr A, 1995, 703:3～15

19 Giddings, JC. J High Resolut Chromatogr, 1987, 10:319

20 Murohy RE, Schure MR, Foley JP. Effect of sampling rate on resolution in comprehensive two-Dimensional liquid chromatography. Anal Chem, 1998, 70:1585～1594

21 Holland LA, Jorgenson JW. Separation of nanoliter samoles of biological amines by a comprehensive two-dimensional microcolumn liquid chromatography system. Anal Chem, 1995, 67: 3275～3283

22 Wagner K, Miliotis T, Marko-Varga G *et al*. An automated on-line multidimentional HPLC system for protein and peptide mapping with integrated sample preparation. Anal Chem, 2002, 74:809～820

23 Optitek GJ, Lewis KC, Jorgenson JW. Comprehensive on-line LC/LC/MS of proteins. Anal Chem, 1997, 69:1518～1524

24 Davis MT, Beierle J, Bures ET *et al*. Automated LC-LC-MS-MS platform using binary ion-exchange and gradient reversed-phase chromatography for improved proteomic analyses. J Chromatogr B, 2001, 752:281～291

25 Vissers JPC, Soest REJ, Chervet J *et al*. Two-dimensional capillary liquid chromatography based on microfractionation. J Microcolumns Sep, 1999, 11:277～286(1999)

26 Opitek G, Jorgenson JW. Two-dimensional SEC/RPLC coupled to mass spectrometry for the analysis of peptides. Anal Chem, 1997, 69:2283～2291

27 Opitek G, Ramirez SM, Jorgenson JW *et al*. Comprehensive two-dimensional high-performance liquid chromatography for the isolation of overexpressed proteins and proteome mapping. Anal Biochem, 1998, 258:349～361

28 Chen Y, Wall D, Lubman DM. Rapid identification and screening of proteins from whole cell lysates human erythroleukemia cells in the liquid phase, using non-porous reversed phase high-performance liquid chromatography separations of proteins followed by multi-assisted laser desorption/ionization mass spectrometry analysis and sequence database searching. Rapid Comm Mass spectrometry, 1998, 12:1994～2003

29 Wall DB, Lubman DM, Flynn SJ. Rapid profiling of induced proteins in bacteria using MALDI-TOF mass spectrometric detection of nonporous RP HPLC-separated whole cell lysates. Anal Chem, 1999, 71:3894～3900

30 Apffel A, Chakel JA, Hancock WS *et al*. Application of multidimensional affinity high-performance liquid chromatograpgy and electrospray ionization liquid chromatography-mass spectrometry to the characterization of glycosylation in single-chain plasminogen activator initial results. J Chromatogr A, 1996, 750:35～42

31 Frank Hsieh YL, Wang H, Elicone C *et al*. Automated analytical system for the examination of protein primary structure. Anal Chem, 1996, 68:455～462

32 Han DK, Eng J, Zhou H *et al*. Quantitative profiling of differentiation-induced microsomal proteins using isotope-coded affinity tags and mass spectrometry. Nat Biotechnol, 2001, 19:946～951

33 Goshe MB, Veenstra TD, Panisko EA *et al*. Phosphoprotein isotope-coded affinity tags: application to the enrichment and identification of low-abundance phosphoproteins. Anal Chem, 2002, 74: 607～616

34 Geng M, Ji J, Regnier FE. Signature-peptide approach to detecting proteins in complex mixtures. J Chromatogr A, 2000, 870: 295～313

35 Moore AW, Jorgenson JW. Comprchensive theree-dimensional separation of peptides using size exclusion chromatography/reversed phase liquid chromatography/optically gated capillary zone electrophoresis. Anal Chem, 1995, 67: 3456～3463

36 MacCoss MJ, McDonald WH, Saraf A *et al*. Shotgun identification of protein modifications from protein complexes and lends tissue. PNAS, 2002, 99:7900～7905

37 McDonald WH, Ohi R, Miyamoto DT *et al*. Comparison of three sirectly coupled HPLC MS/MS strategies for identification of proteins from complex mixtures: single-dimension LC-MS/MS, 2-phase MudPIT, and 3-phase MudPIT. Int

J Mass Spectrom, 2002, 219:245～251

38 Link AJ, Eng J, Schieltz DM *et al*. Direct analysis of protein complexes using mass spectrometry. Nat Biotechnol, 1999, 17:676～682

39 Washburn MP, Wolter D, Yates JR. Large-scale analysis of the yeast proteome by multidimensional protein identification technology. Nat Biotechnol, 2001, 19:242～247

40 Wolters DA, Washburn MP, Yates JR. An automated multidimensional protein identification technology for shotgun proteomics. Anal Chem, 2001, 73: 5683～5690

41 Washburn MP, Ulaszek R, Deciu C *et al*. Analysis of quantitative proteomic data generated via multidimensional protein identification technology. Anal Chem, 2002, 74:1650～1657

42 Issaq HJ, Conrads TP. Methods for fractionation, separation and profiling of proteins and peptides. Electrophoresis, 2002, 23:3048～3061

43 vensson H. Isoelectric fractionation, analysis, and characterization of ampholytes in natural pH gradients. I. The differential equation of solute concentrations at a steady state and its solution for simple cases. Acta Chem Scand, 1961, 15: 325～341

44 vensson H. Isoelectric fractionation, analysis, and characterization of ampholytes in natural pH gradients. II. Buffering capacity and conductance of isoionic ampholytes. Acta Chem Scand, 1961, 16:456～466

45 Svensson H. Prot Biol Fluids, 1967, 15:515～522

46 Tattanahalli L, Sharma NB, Trotta PP. Application of recycling isoelectric focusing for purification of recombinant leukocyte interferons. Electrophoresis, 1986, 7:552～557

47 Wall DB, Kachman MT, Lubman DM. Isoelectric focusing nonporous RP HPLC: a two dimensional liquid-Phase separation method for mapping of cellular proteins with identification using MALDI-TOF mass spectrometry. Anal Chem. 2000, 72:1099～1111

48 Kachman MT, Lubman DM. A 2-D Liquid separations/mass mapping method for interlysate comparison of ovarian cancers. Anal Chem, 2002, 74:1779～1791

49 Herbert B, Righetti PG. A turning point in proteome analysis: sample pre-fractionation via multicompartment electrolyzers with isoelectric membranes. Electrophoresis, 2000, 21: 3639～3648

50 Bier M, Egen NB, Allgyer TT. Peptides: Structure and Biological Function, Pierce Chemical Co. Rockford, IL 1979, 35～48

51 Haglund H, Westerfeld J G, Ball J T. eds., Electrofocus 78, Elsevier, Amsterdam 1978, 35～48

52 Bier M, Egen NB, Mosher RA *et al*. In: Rindone G. ed. Materials processing in the microgravity environment of space, New York: Elsevier Science, 1982. 261～266

53 Bier M. Recycling isoelectric focusing and isotachophoresis. Electrophoresis, 1998, 19:1057～1063

54 Hanning K. New aspects in preparative and analytical continuous free-flow cell electrophoresis. Electrophoresis, 1982, 3:235～243

55 Janson JC, Ryden L. eds. Protein purification, 2nd ed, New York: Wiley-VCH, 1988

56 Hille MJ, Freed AL. Possibilities to improve automation, speed and precision of proteome analysis: a comparison of two-dimensional electrophoresis and alternatives., Electrophoresis, 2001, 22:4035～4052

57 McDonald WH, Yates JR 3rd., Shotgun proteomics and biomarker discovery. Dis Markers, 2002, 18(2):99～105

58 Wu CC, MacCoss MJ. Shotgun proteomics: tools for the analysis of complex biological systems. Curr Opin Mol Ther, 2002, 4(3):242～250

59 萨姆布鲁克 J,拉塞尔 D W 著.黄培堂等译.分子克隆实验指南(上下册)第 3 版,北京:科学出版社,2002. 1604

60 Nagabhushan TL, Sharma B, Trotta PP. Application of recycling isoelectric focusing for purification of recombinant human leukocyte interferons., Electrophoresis, 1986, 7:552～557

61 Wall DB, Maureen T *et al*. Isoelectric focusing nonporous RP HPLC: a two-dimensional liquid-phase separation method for mapping of cellular proteins with identification using MALDI-TOF mass spectrometry. Anal Chem, 2000, 72: 1099～1111

62 Kachman MT, Haixing Wang *et al*. A 2-D liquid separations/mass mapping method for interlysate comparison of ovarian cancers. Anal Che, 2002, 74:1779～1791

63 Davidsson P, Westman A, Puchades M. Characterization of proteins from human cerebrospinal fluid by a combination of preparative two-dimensional liquid-phase electrophoresis and matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. Anal Chem, 1999, 71:642～647

64 Herbert B, Righetti PG, A turning point in proteom analysis: Sample prefractionation via multicompartment electrolyzers with isoelectric membranes. Electrophoresis, 2000, 21:3639～3648

65 Bier M, Twitty GE Sloan JE. Recycling isoelectric focusing and isotachophoresis., J Chromatogr, 1989, 470:369～376

66 He Y, Zhong W, Yeung ES. Multiplexed on-column protein digestion and capillary electrophoresis for high-throughput comprehensive peptide mapping. J Chromatogr B,2002, 782:331～341

67 Bindila L, Zamfir A, Peter-Katalinic J. Characterization of peptides by capillary zone electrophoresis and electrospray ionization quadrupole time-of-flight tandem mass spectrometry. J Sep Sci,2002, 25:1101～1111

68 Wetterhall M, Palmblad M, Hakansson P *et al*. Rapid analysis of tryptically digested cerebrospinal fluid using capillary elecrophoresis-electrospray ionization-fourier transform ion cyclotron resonance-mass spectrometry. J Proteome Res, 2002, 1:361～366

69 Severs JC, Smith RD. Electrospray ionization mass spectrometry, fundamentals, instrumentation, and applications. New York: Wiley, 1997. 343～382

70 Covey TR, Pinto D. Applied electrospray mass spectrometry. New York: Marcel Dekker, 2002. 5～148

71 Shen Y, Berger SJ, Anderson GA *et al*. High-efficiency capillary isoelectric focusing of peptides. Anal Chem,2000, 72:2154～2159

72 Shen Y, Xiang F, Veenstra TD *et al*. High-resolution capillary isoelectric focusing of complex protein mixtures from lysates of microorganisms. Anal Chem,1999, 71:5348～5353

73 Manabe T, Miyamoto H, Inoue K *et al*. Separation of human cerebrospinal fluid proteins by capillary isoelectric focusing in the absence of denaturing agents. Electrophoresis,1999, 20:3677～3683

74 Shen Y, Smith RD. High-resolution capillary isoelectric focusing of proteins using highly hydrophilic-substituted cellulose-coated capillaries. J Microcol Sep,2000, 12:135～141

75 Martinovic S, Berger SJ, Pasa-Tolic L *et al*. Separation and detection of Intact noncovalent protein complexes from mixtures by on-line capillary isoelectric focusing-mass spectrometry. Anal Chem,2000, 72:5356～5360

76 Armstrong DW, Schulte G, Schneiderheinze JM *et al*. Separating microbes in the manner of molecules. 1. Capillary electrokinetic approaches. Anal Chem,1999,7871:5465～5469

77 Shen Y, Beger S, Smith RD. Capillary isoelectric focusing of yeast cells. Anal Chem,2000, 72:4603～4607

78 Sheng L, Pawliszyn J. Comprehensive two dimensional separation based on coupling micellar electrokinetic chromatography with capillary isoelectric focusing. Analyst,2002,127: 1159～1163

79 Mohan D, Lee CS. On-line coupling of capillary isoelectric focusing with transient isotachophoresis-zone electrophoresis: A two-dimensional separation system for proteomics. Electrophoresis,2002, 23: 3160～3167

80 Lemmo AV, Jorgenson JW. Transverse flow gating interface for the coupling of microcolumn LC with CZE in a comprehensive two-dimensional system. Anal Chem,1993, 65:1576～1581

81 Lemmo AV, Jorgenson JW. Two-dimensional protein separation by microcolumn size-exclusion chromatography-capillary zone electrophoresis,J chromatogr,1993, 633: 213～220

82 Moore AWJr, Jorgenson JW. Rapid comprehensive two-dimensional separations of peptides via RPLC-optically gated capillary zone electrophoresis. Anal Chem,1995, 67: 3448～3455

83 Lewis KC, Opiteck GL, Jorgenson JW *et al*. Comprehensive on-line RPLC-CZE-MS of peptides. J Am Soc Mass Spectrom,1997, 8: 495～500

84 Hooker TF, Jorgenson JW. A Transparent flow gating interface for the coupling of microcolumn LC with CZE in a comprehensive two-dimensional system. Anal Chem,1997, 69:4134～4142

85 Chen J, Lee CS, Shen Y *et al*. Integration of capillary isoelectric focusing with capillary reversed-phase liquid chromatography for two-dimensional proteomics separation. Electrophoresis, 2002, 23: 3143～3148

86 Cao P, Stults JT. Phosphopeptide analysis by on-line immobilized metal-ion affinity chromatography-capillary electrophoresis-electrospray ionization mass spectrometry. J Chromatogr A, 1999, 853:225～235

87 Cao P, Stults JT. Mapping the phosphorylation sites of proteins using on-line immobilized metal affinity chromatography/capillary electrophoresis/electrospray ionization multiple stage tandem mass spectrometry. Rapid Commun Mass Spectrom, 2000, 14:1600～1606

88 Figeys D, Gygi SP, Zhang Y *et al*. Electrophoresis combined with novel mass spectrometry technigues: Powerful tools for the analysis of proteins and proteomes. Electrophoresis, 1998, 19:1811～1818

89 Issaq HJ, Chan KC, Janini GM *et al*. A simple two-dimensional high performance liquid chromatography/high performance capillary electrophoresis set-up for the separation of complex mixtures. Electrophoreisi, 1999, 20:1533～1537

90 Tong W, Link A, Eng JK *et al*. Identification ofproteins in complexes by solid-phase microextraction/multistep elution/capillary electrophoresis/tandem mass spectrometry. Anal Chem, 1999, 71: 2270～2278

91 Hofmann O, Che D, Cruickshank KA *et al*. Adaptation of capillary isoelectric focusing to microchannels on a glass chip. Anal Chem, 1999, 71:678～686

92 Li JJ, Thibault P, Bings NH *et al*. Integration of microfabricated devices to capillary electrophoresis-electrospray mass spectrometry using a low dead volume connection: application to rapid analyses of proteolytic digests. Anal Chem, 1999, 71: 3036～3045

93 Chan JH, Timperman AT, Qin D *et al*. Microfabricated polymer devices for automated sample delivery of peptides for analysis by electrospray ionization tandem mass spectrometry. Anal Chem, 1999, 71: 4437～4444

94 Zhang BL, Foret F, Karger BL. A microdevice with integrated liquid junction for facile peptide and protein analysis by capillary electrophoresis/electrospray mass spectrometry. Anal Chem, 2000, 72:1015～1022

95 Gottschlich N, Culbertson CT, Mcknight TE *et al*. Integrated microchip-device for the digestion, separation and post-column labeling of proteins and peptides. J Chromatogr B, 2000, 745:243～249

96 Rocklin RD, Ramsey RS, Ramsy JM. A microfabricated fluidic device for performing two-dimensional liquid-phase separations. Anal Chem, 2000, 72:5244～5249

97 Jacobson SC, Ramsey JD, Culbertson CT *et al*. Proceedings of micro TAS 2002 symposium. Dordrecht. (The Netherlands). Kluwer Academic Publishers, 2002, 2:609

98 Wang Z, Swinney K, Bornhop DJ. Attomole sensitivity for unlabeled proteins and polypeptides with on-chip capillary electrophoresis and universal detection by interferometric backscatter. Electrophoresis, 2003, 24:865～873

99 Walhagen K, Unger K, Hearn M. Capillary electrocheomatography analysis of hormonal cyclic and linear peptides. Anal Chem, 2001, 73:4924～4936

100 Walhagen K, Unger K, Keah H *et al*. Application of CEC procedures for the analysis of synthetic peptides: characterization of linear immunogenic peptides that mimic a HIV-1 gp120 epitope. J Pept Res, 2002, 59:159～173

101 Walhagen K, Unger K, Keah H *et al*. Influence of temperature on the behavior of small linear peptides in capillary electrochromatography. J Chromatogr A, 2000, 893: 401～409

102 Ye M, Zou H, Liu Z *et al*. Separation of peptides by strong cation-exchange capillary electrocheomatography. J Chromatogr A, 2000, 869:385～394

103 Ye M, Zou H, Liu Z *et al*. Capillary electrochromatography with a silica column dynamically modified cationic surfactant. J Chromatogr A, 1999, 855:137～145

104 Matyska M, Pesek J, Boysen R *et al*. Electrochromatographic characterization of etched chemically-modified capillaries with small synthetic peptides. J Chromatogr A, 2001, 924:211～221

105 Huang X, Zhang J, Horvath C. Capillary electrochromatography of proteins and peptides with porous-layer open-tubular columns. J Chromatogr A, 1999, 858:91～101

106 Xu W, Regnier F Electrokinetically-driven cation-exchange chromatography of proteins and its comparison with pressure-driven high-performance liquid chromatography. J Chromatogr A, 1999, 853:243～256

107 Palm A, Novotny MV. Macroporous polyacrylamide/poly(ethylene glycol) matrixes as stationary phases in capillary electrochromatography. Anal Chem, 1997, 69: 4499～4507

108 Erison C, Hjerten S. Reversed-phase electrochromatography of proteins on modified continuous beds using normal-flow and counterflow gradients. Theoretical and practical considerations. Anal Chem, 1999, 71:1621～1627

109 Gusev I, Huang X, Horvath C. Capillary columns with in situ formed porous monolithic packing for micro high-performance liquid chromatography and capillary electrochromatography. J Chromatogr A, 1999, 855:273～290

110 Zhang S, Huang X, Zhang J *et al*. Capillary electrochromatography of proteins and peptides with a cationic acrylic monolith. J Chromatogr A, 2000, 887:465～477

111 Wu R, Zou H, Fu H *et al*. Separation of peptides on mixed mode of reversed-phase and ion-exchange capillary electrochromatography with a monolith column. Electrophoresis, 2002, 23:1239～1245

112 Wu R, Zou H, Ye M *et al*. Capillary electrochromatography for separation of peptides driven with electrophoretic mobility on monolithic column. Anal Chem, 2001, 73: 4918～4923

113 He B, Tait N, Regnier F. Fabrication of nanocolumns for liquid chromatography. Anal Chem, 1998, 70:3790～3797

114 He B, Ji J, Regnier F. Capillary electrochromatography of peptides in a microfabricated system. J Chromatogr A, 1999, 853:257～262

115 Slentz BE, Penner N, Lugowska E *et al*. Nanoliter capillaru electrochromatography columns based on collocated monolithic support structures molded in poly(dimethylsiloxane). Electrophoresis, 2001, 22:3736～3743

116 Slentz BE, Penner NA, Regnier FE. Protein proteolusis and the multi-dimensional electrochromatographic separation of histidine-containing peptide fragments on a chip. J Chromatogr A, 2003, 984:97～107

第二十一章　蛋白质相互作用和蛋白质芯片

第一节　蛋白质相互作用

一、概　　况

随着人类基因组测序工作的完成,生命科学进入了后基因组和蛋白质组时代。因此,作为其中一部分的蛋白质相互作用的研究显得愈来愈重要。生命活动的过程与蛋白质之间的相互作用密不可分。例如,DNA 合成、基因转录激活、蛋白质翻译、细胞周期调控、信号传导等重要的生命过程均涉及到蛋白质复合体的作用。我们以 Wnt 信号通路为例来说明[1]。Wnt 通路是一条保守性很强的信号通路,该通路调节控制了许多生命过程,如细胞形态与功能的分化与维持,免疫,应激,细胞癌变与细胞凋亡,等等。Wnt 信号通路的作用分子包括:Wnt 蛋白家族成员、卷曲(frizzled)蛋白、Dishevelled 蛋白、β-联蛋白(β-catenin)、轴蛋白(axin)、结肠癌抑制因子(APC)、糖原合酶激酶(GSK3β)、β-TrCP 蛋白、淋巴增强因子(LEF)/T 细胞因子(TCF)等。当没有 Wnt 信号时,GSK3β、APC、axin 组成破坏复合体,使 β-联蛋白被磷酸化,最终泛肽化而降解。细胞核内,转录抑制因子 Groucho 家族成员与转录因子 TCF 形成复合物,通过 HMG 框结合在靶基因上,抑制靶基因的转录。当有 Wnt 信号传入时,通路中的下游分子 Dsh 抑制了破坏复合体的作用,β-联蛋白在胞质中积累进入核内,TCF 与入核的 β-联蛋白结合,导致其与 Groucho 的结合下降,从而去除抑制作用,激活了靶基因的转录(图 21.1)。

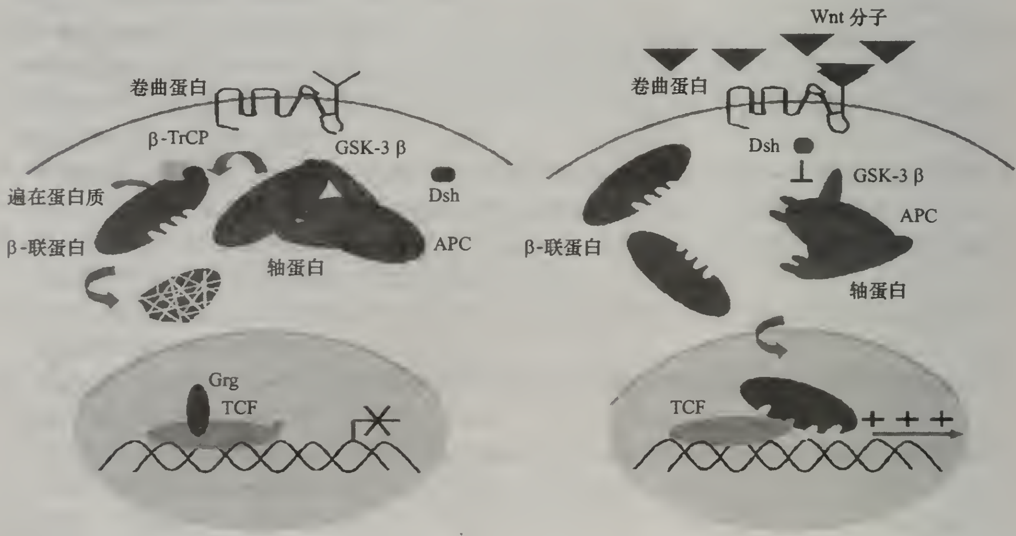


图 21.1　Wnt 信号通路

蛋白质相互作用包括三方面。一是多亚基蛋白质的形成，即分离纯化后可得到两个或多个不同组分。例如血红素、色氨酸合成酶、*E. coli* 的 DNA 合成复合体等。二是多成分的蛋白质复合物，如核孔复合体、剪接体、纺锤体等。三为瞬时的蛋白质相互作用，控制着一些重要的细胞内活动。所有的蛋白质修饰过程都需要这类相互作用。如蛋白激酶、蛋白磷酸酶、糖基转移酶、乙酰基转移酶、蛋白酶等酶与各自蛋白质底物之间的相互作用。这类相互作用几乎参与调节了细胞内基本生命活动的所有形式，如细胞生长、细胞周期、代谢途径、信号传导等。除了蛋白质修饰以外，其他过程如转录复合体与特定启动子的结合、蛋白质的跨膜运输、新生肽链的折叠也包含了瞬时的蛋白质相互作用。

二、蛋白质相互作用研究策略和方法

（一）生化方法

1. 蛋白质亲和层析

在一定的条件下，某种蛋白质可以共价连接在基质如琼脂糖（Sepharose）上，用以结合抽提物中能与该种蛋白质结合的配体蛋白质。抽提物过柱后，用低盐溶液洗脱下未结合的物质，然后用高盐溶液或 SDS 洗脱下结合在柱子上的蛋白质（图 21.2）。有时污染物的存在会影响相互作用的结果，因此蛋白质纯化是蛋白质亲和层析成功的前提。获得纯化蛋白质的一个简单方法是利用融合蛋白，目前常用的为谷胱甘肽 S 转移酶（glutathione S-transferase，GST），可以在谷胱甘肽-琼脂糖（agarose）柱上纯化[2]。其他的还有 *Staphylococcus* 蛋白 A 可在含 IgG 的柱上纯化；含少数组氨酸的肽段在镍柱上纯化；麦芽糖结合蛋白可在含直链淀粉的柱子上纯化。在亲和柱上，可以利用纯化的融合蛋白以两种方式检测相互作用[3~7]。一种是将融合蛋白共价连接在树脂上，让抽提物过柱。二是把融合蛋白非共价地结合在珠子（bead）上，使抽提物和珠子混合在一起。

利用蛋白质亲和层析来检测蛋白质相互作用有以下优点。①灵敏度高。在高浓度固定蛋白质的条件下，可以检测微弱的相互作用（结合常数为 10^{-5}mol/L）[8]，而生理上相关的最微弱的相互作用在 10^{-3}mol/L 范围内。②可以同时检测抽提物中的所有蛋白质，即所有蛋白质与固定蛋白质的结合机会均等。③通过构建突变体，可以检测与特异性相互作用有关的蛋白质结构域或关键的氨基酸残基。④可以检测多亚基蛋白质之间的相互作用。

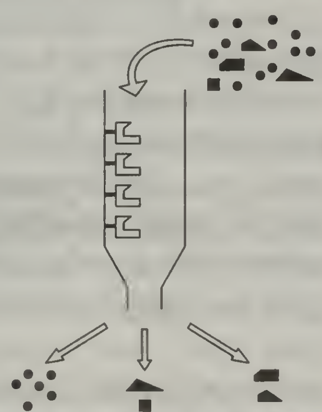


图 21.2　蛋白质亲和层析图

虽然蛋白质亲和层析技术能有效地检测蛋白质相互作用，但是也存在着一些假阳性的结果。原因可能有以下几点。①电荷间的相互作用使蛋白质与检测蛋白质结合在一起，可以利用带相同离子电荷的对照柱来消除。②通过 B 蛋白，A 蛋白与检测蛋白质间接结合。③两种蛋白质的细胞定位不同，但是在人为条件下，却能以高特异性结合。一个著名的例子就是肌动蛋白与 DNase I 之间的高亲和力[9]。

2. 亲和印迹

蛋白质用 PAGE 胶分离后，转移到硝酸纤维素膜上，然后用特异性的蛋白质、肽段或配体来检测膜上的蛋白质。该技术与 Western 印迹有些类似，只不过 Western 是利用抗体作为探针。复杂的蛋白质混合物，如全细胞裂解液，通过该技术可以直接分析，而不需要纯化。因此，该方法特别适合于分析膜蛋白，如细胞表面受体[10]。在凝胶电泳之前，可以先分离细胞裂解物以提高该方法的灵敏度，使得低丰度蛋白质的相互作用也可以被检测到。

该方法涉及到以下几个关键：膜上蛋白质的生物活性、蛋白质探针的制备、检测方法。通常混合物在有 SDS 存在的情况下分离，而在印迹时需要去除变性剂，使变性蛋白质全部或部分复性。如果蛋白质在变性后无法复性，那就需要一个非变性的胶分离系统。蛋白质探针的制备有多种方法。探针可以用放射性标记，也可以用生物素标记，或者使用亲和标签等方法。

3. 免疫共沉淀

免疫共沉淀是一种经典的检测蛋白质相互作用的方法，其实验过程比较简单。裂解细胞后，加入抗体，抗原被沉淀下来后洗涤，去除非特异性结合，再分析结合复合体。抗体可以是单克隆，也可以是多克隆。被分析的蛋白质可以加上一个抗原决定簇的标签，如 12CA5、c-Myc，便于用特异性的抗体检测。

免疫共沉淀实验的真实性需要以下几条标准来验证。①抗原的沉淀是由于本身抗体的结合引起的，而不是制备中被污染的抗体。单克隆抗体不存在这个问题。在使用多克隆抗体时，需要对该抗体进行预处理，即将抗体与不包含抗原的抽提物混合在一起，去除与抗体结合的污染物。②只有在抗原存在的情况下，加入抗体后，才能检测到沉淀蛋白质，即抗体本身不识别沉淀蛋白质。③这种相互作用是否直接或是通过第三个蛋白质引起的。④相互作用发生在细胞内，而不是细胞裂解的结果。

免疫共沉淀有以下优点。①与蛋白质亲和层析一样，它检测的是细胞裂解原液中所有蛋白质与检测蛋白质的相互作用。②抗原和抗体的相互作用是在生理浓度的条件下被检测的，因此过量表达所带来的假阳性结果可以被排除。③可以检测到在体内形成的天然复合体，如果该复合体无法在体外条件形成。④天然状态的蛋白质如果经过翻译后修饰，依赖于或不依赖于修饰的蛋白质相互作用可以被检测到。虽然该方法应用广泛，但是也存在着一些缺点。①有时免疫共沉淀的蛋白质并非直接相互作用。例如，E1A 与 p107 可相互作用[11]，p107 与 cyclinA 可相互作用[12,13]，E1A 与 cyclinA 能同时被免疫沉淀[14]。②与其他方法（如蛋白亲和层析）相比，免疫共沉淀的灵敏度不高，这是由于抗原浓度低于亲和层析中的蛋白质浓度造成的。

4. 化学交联法

使用一种交联试剂，并且在交联试剂的光激活部分带有标记，将该试剂耦联到一蛋白质上，耦联体与抽提物温育在一起。如果抽提物中的 A 蛋白能与耦联蛋白质发生相互作用，光激活后交联剂的一部分就结合到 A 蛋白上。加入还原剂，切割交联试剂，使 A 蛋白

上带有交联剂上有标记的部分,然后通过凝胶电泳等方法分析该蛋白质(图 21.3)。常用的交联剂有: Denny-Jaffee 试剂[15], ^{125} *I*-{*S*-[*N*-(3-iodo-4-azidosalicyl) cysteaminyl]-2-thiopyridine}即 IAC[16]等。运用此方法,国外学者在顶体完整的小鼠精子中鉴别了一个与 ZP3 结合的特异的分子质量为 56 000Da 的蛋白质[17],并且用亲和层析获得同样相对分子质量的蛋白质。

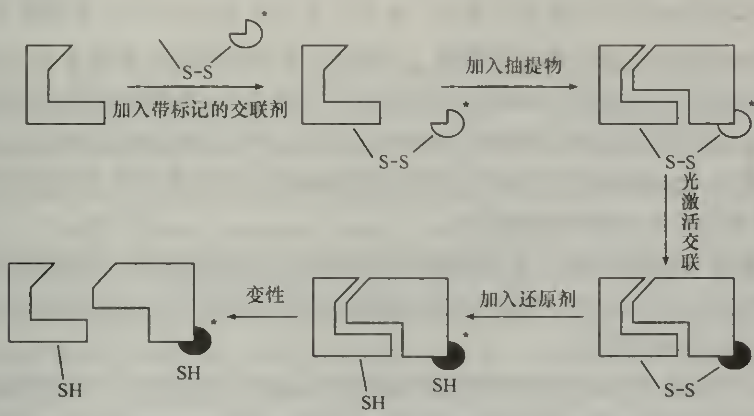


图 21.3　化学交联法示意图

利用交联技术检测蛋白质相互作用的优点在于:①它能检测到微弱的相互作用[18];②能检测到动态过程不同阶段与不同蛋白质的瞬时相互作用,如糖基化过程;③通过能穿膜的交联试剂,交联可在体内发生[18]。但是该技术的主要缺陷是没有选择性,交联试剂可与近范围内的蛋白质发生反应。所以,有时所检测到的蛋白质相互作用可能是个假象,需要通过其他方法来验证。

5. 荧光共振能量转移(fluorescence resonance energy transfer, FRET)**法**

FRET 是一个无辐射、量子级能量转移现象,它指的是当一个荧光分子(即供体)受到激发时,能量向邻近的另一荧光基团(即受体)转移的过程。但仅当受体与供体间的距离小于 10 nm 且受体的吸收光谱和激发光谱有重叠时,FRET 才能发生。并且随着光谱重叠的增加,FRET 的效率将增高,同时 FRET 可能发生的距离也将增加[19]。

荧光基团距离和空间取向的变化,将引起 FRET 效率的改变。将 GFP 的蓝移突变型与红移突变型和一段特异的多肽链连接起来,通过观察两个波长发射光比值的变化,来记录 FRET 效率的变化[20]。单细胞的 FRET 效率可以用 FRET 显微镜得到,而多细胞的平均 FRET 可以用荧光光度计或流式细胞仪得到。

使用 FRET 可用来检测蛋白质之间的相互作用,将两蛋白质分别使用 CFP/YFP(或 BFP/GFP)标记,当两者在同一细胞中表达时,只要检测到 FRET 发生,则证明两蛋白质间距离小于 10nm,存在相互作用。此方法已经成功地验证了 CFP-钙调蛋白与 M13-YFP 融合蛋白间、FHL3 与 FHL2、EGF 受体与 Grb2 等蛋白质之间相互作用[19~21]。

GFP-FRET 将活体高分辨率观测和细胞、亚细胞定位相结合,使得用显微镜观察纳米水平的距离变化成为可能,为人们在活体中观察生物大分子的相互作用提供了一个有效的方法。但是,在活体 GFP-FRET 的研究中,常常存在着信号太弱、细胞自发荧光等不

利因素影响实验结果。

6. 生物传感器技术,即 BIA (biomolecular interaction analysis)-core

生物传感器的工作原理是基于表面胞质共振(surface plasmon resonance, SPR),即配体和分子结合时会引起作用表面(平面或近似平面)折射率的变化。这种光信号可被光感受器接收并转换成电信号传送到计算机,再由计算机还原为模拟的生物信号。BIA-core 以共振单位(resonance unit, RU)来衡量,1 000RU 相当于每平方毫米 1 纳克的质量[22]。

采用生物传感器检测芯片上配体与待检分子的结合时,需要考虑很多参数的影响,如原料输送速率、配体种植密度、待检分子浓度,以及这种结合本身的亲和力大小等。原料输送速率是指分子从输送流层投递到芯片表面的速率。这些参数可直接或间接地被实时测量,由计算机推算出初始结合速率。

当原料在芯片表面流动时,在原料输送到和未输送到的界面处,原料输送流层和芯片表面的分子浓度是有差异的。现代生物传感器采用双监测模型,应用两套实时监测设备,同时对二者进行分别监测[22]。

BIA-core 除具有现代生物传感器的基本特点外,最大的特色在于其附带的分析软件。它内置若干描述蛋白质相互作用的方程模型,在模型的基础上根据采集的数据给出结合速率对各参数的偏微分方程,计算出初始结合率。采用人机对话框,可根据要求作图和给出数据。同时生物传感器也可与质谱联用,直接监测芯片上配体所捕获的分析物分子的相对分子质量(图 21.4)[23]。

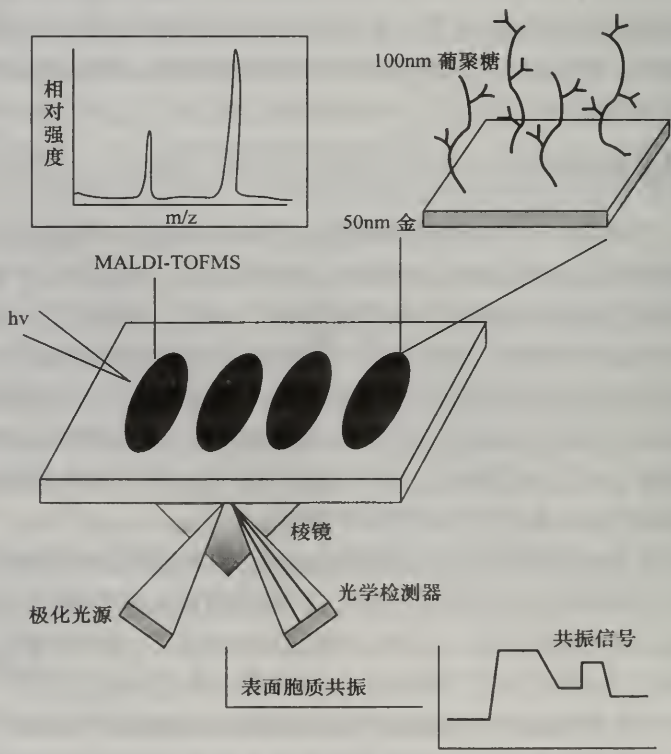


图 21.4 BIA-core 工作原理示意图

目前,BIA-core 被广泛应用于检测抗原-抗体相互作用[24~28]、受体-配体相互作用[29~33]、蛋白质-核酸相互作用[34,35]、蛋白质-小分子物质相互作用[36]等。同时 BIA-core 也可应用于监测配体的纯化,如对人的类 EPH 激酶配体的纯化[37]。

(二) 分子生物学方法

1. 酵母双杂交

酵母双杂交系统于 1989 年由 Fields 和 Song 创立,是用来研究蛋白质与蛋白质之间相互作用的简便而有效的方法,又称为"相互作用陷阱"[38]。双杂交系统通常在酵母细胞中构建,首先需要构建两种杂合反式作用因子,将蛋白质 X 与报道基因转录因子的特异 BD(DNA binding domain,DNA 结合结构域,如 Gal4-BD, LexA-BD)融合,成为"诱饵"蛋白。蛋白质 Y 与特异的 AD(activation domain,转录激活结构域,如 Gal4-AD, LexA-AD)融合为"猎物"蛋白。当编码两种结构域的基因在细胞核内同时表达时,若蛋白质 X 和 Y 之间存在相互作用,就会使 AD 和 BD 两种结构域的上游活化序列相互接近,进而激活转录过程,使报道基因得到表达。目前应用最多的是 *E. coli* 的 *LacZ* 基因,它表达后能在 X-Gal 的培养基上产生蓝色菌落,阳性结果易于判断。除此之外,营养缺陷型筛选也应用较多:通过在缺乏相应氨基酸(如组氨酸、亮氨酸)的培养基上菌落生长与否来判断 *His*3、*Leu*2 等报道基因是否表达。现在多提倡两种报道基因同时应用,即颜色筛选与营养筛选同时进行。双杂交系统普遍构建于酵母细胞中的原因,是因为酵母细胞便于转化,结果易于筛选;而且内源性酵母蛋白质极少与哺乳细胞靶蛋白质结合,避免了对文库编码蛋白质的干扰。

自创立以来,酵母双杂交体系迅速发展成为一种常规的分子生物学技术,并得到了广泛的应用。近年来,国外学者在原有的酵母双杂交的基础上,发展了反向双杂交体系(reverse two-hybrid system)和三杂交体系(three-hybrid system)。

(1) *反向双杂交体系*

反向双杂交体系的核心在于构建一种反向筛选的报道基因,即蛋白质间特异相互作用所激活的报道基因表达能阻碍转化体生长,从而形成发现蛋白质间结合解离的选择环境。已知酵母 *URA3* 基因产物一方面为合成尿嘧啶所必需,同时又能催化 5-flouroorotic acid(5-FOA)转变为一种有毒物质[39](图 21.5 A)。在 Vidal 等构建的酵母菌株中,*URA3* 基因的表达受到含 Gal4 结合位点的启动子的调控[40]。利用尿嘧啶的缺陷培养可以筛选出与"诱饵"蛋白相互作用的蛋白质(阳性筛选),而在添加了 5-FOA 的培养基中,存在蛋白质相互作用的菌株却会因毒素作用而死亡(阴性筛选),只有发生了蛋白质间解离或部分解离的转化株才可以表现出抗性而存活。另一反向筛选系统为分离杂交体系(split-hybrid system)[41](图 21.5 B),在此体系中存在着两个报道基因。第一个报道基因包含着 *E. coli* tet 抑制子(TetR)的编码序列,第二个报道基因为 *His*3,启动子区插入了 tet 操纵子序列。双杂交产生的相互作用激活 Tet 阻遏蛋白,从而抑制阳性筛选标志 *His*3 的表达,此时野生型相互作用使宿主细胞表现为组氨酸营养缺陷型。而只有那些作用蛋白质间已发生解离或部分解离的转化株可以在组氨酸缺陷的培养基上生长。

反向双杂交体系作为双杂交体系的补充,在研究蛋白质相互作用中发挥着日益重要的作用。①在筛选蛋白质突变体时,反向双杂交体系能从大量突变体中直接筛选出导致

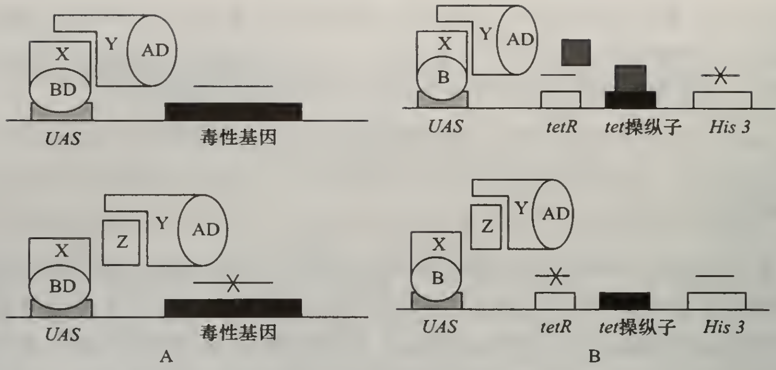


图 21.5　反向双杂交体系

蛋白质间相互作用减弱或消失的突变位点,而传统的利用 *Lac* 报道基因来筛选的工作量大,且带有盲目性。②反向双杂交能够发现可导致已知蛋白质间特异相互作用发生解离的肽类或其他小分子物质。③传统双杂交体系的假阳性问题来源于文库中的许多片段具有自动激活报道基因表达的活性。如果能够应用反向双杂交对文库进行"预清除"[42],则有可能大大减轻以后的假阳性鉴定的负担。

(2) 三杂交体系

三杂交体系是在双杂交体系上发展起来的另一分支,主要用于研究多蛋白质复合体之间的相互作用[43~46]。融合蛋白 BD-X 和 AD-Y 在酵母细胞核内表达时,同时有第三种蛋白质 Z 表达。X 和 Y 不能直接发生相互作用,必须通过桥梁——蛋白质 Z,X 和 Y 才能发生相互作用,从而激活下游报道基因的转录(图 21.6)。

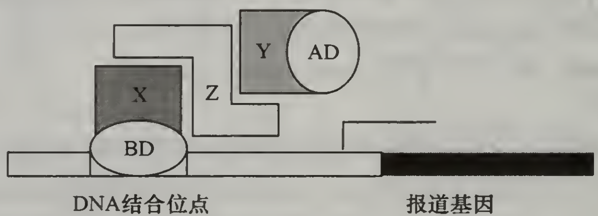


图 21.6　三杂交体系

由于在单个酵母细胞中,传统的双杂交体系只能研究两个蛋白质之间的相互作用,所以,在研究蛋白质复合体时,由于没有复合体中其他蛋白质的作用,多肽链可能会暴露出生理状态下为其他蛋白质所覆盖的某些位点,因此造成假阳性或假阴性的结果。而在三杂交体系中,加入表达受到调控的第三种蛋白质可以使蛋白质复合体的相互作用更符合生理状态。因此,三杂交体系是对双杂交系统的一个有益补充。以此为基础,甚至可以再构建四杂交体系和反向三杂交体系,以研究更广泛的蛋白质间相互作用[43]。

同以往研究蛋白质相互作用的实验手段相比,双杂交系统具有独特的优势。融合体蛋白质之间的相互作用在真核酵母细胞内进行,蛋白质有可能保持天然的折叠状态,类似其在体内生理状态下的情况,这是其他体外生化检验方法所缺乏的。因此与后者相比,它所证实的蛋白质间的相互作用将更接近于体内的真实水平。双杂交系统最大的应用价值在于它可以筛选 cDNA 文库中编码与已知蛋白质特异作用的成分。这种筛选过程不但揭

示了许多已知蛋白质之间的未知作用，而且有助于发现许多新基因。双杂交系统检测到的相互作用在体内发生，不需要额外的纯化步骤。此外，它能检测到短暂的相互作用，目前被广泛应用于信号通路的研究。

但是，酵母双杂交系统也存在着一些局限性，包括：①它只研究在核内表达的蛋白质的相互作用，融合蛋白的构建可能会影响到蛋白质本身的活性或折叠状态；②如果特定的翻译后修饰在酵母细胞中不发生，则不能检测到修饰后蛋白质的相互作用；③许多蛋白质与 BD 融合后，具有自激活效应，导致假阳性结果[47]。

2. 噬菌体展示

噬菌体能在其表面表达一种融合蛋白，该融合蛋白包含外源肽段，可与抗体结合[48]。通过免疫亲和纯化，“融合型”噬菌体得到富集。将文库基因构建到噬菌体中，然后筛选出与特异抗体相互作用的蛋白质，这一过程被称为“淘选”。抗体通常固定在塑料盘上，不能结合的噬菌体被洗脱掉，而结合的噬菌体被保留下来，感染 *E. coli*。经过几个周期后，可以富集结合紧密的肽段的 DNA 序列(图 21.7)。

丝状噬菌体如 M13、fd、f1 在其表面大约有 5 个拷贝基因编码的被膜蛋白，因此异源的 DNA 序列可以插入到该基因中，得到多拷贝的融合蛋白，被称为“多价展示”。而另一编码被膜蛋白的基因可以产生 2 700 个融合蛋白的拷贝[49]。通常情况下，多价展示只局限于小肽段，因为插入大片段会影响被膜蛋白的功能和噬菌体的感染力。

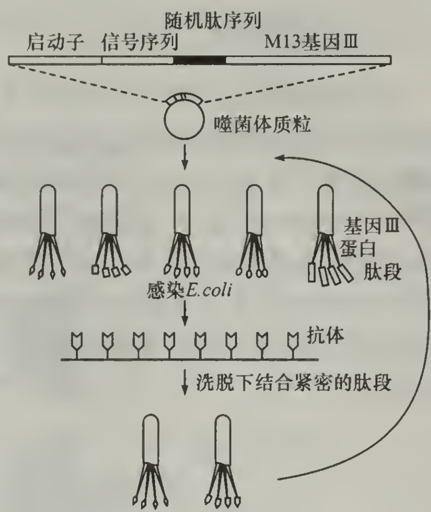


图 21.7　噬菌体展示

在研究蛋白质相互作用方面，噬菌体展示有以下优点。①高度的选择性。每一周期中筛选出的噬菌体感染 *E. coli* 后，该噬菌体可得到 1 000 倍甚至更多的扩增。几次循环后，可得到相互作用的序列。②可以构建大的噬菌体文库(10^9～10^{10})，亲和纯化可在高浓度(＞10^3噬菌体/ml)下进行[50]。③通过改变洗脱步骤的严格性，来评价融合蛋白结合的特异性。

但是该方法也存在着一些缺点：①多价展示中对于肽段大小的限制；②蛋白质必须要有能被 *E. coli* 分泌的特性；③因为是体外实验，蛋白质在宿主细菌中有时无法正确折叠或修饰；④构建的融合蛋白会影响到蛋白质本身的活性。

（三）遗传学方法

1. 基因外抑制子

抑制子的突变能部分或全部逆转原有突变，使之成为正常表型。如图 21.8 所示，野生型蛋白质 X、Y 能发生相互作用，当 X 突变时，X^* 与 Y 不发生作用，而当 Y 也突变后，突变体 X^* 和 Y^* 又能重新形成二聚体，恢复原有的相互作用。

该方法应用于线虫(*Caenorhabditis elegans*)的研究，因其个体数量巨大，可以进行大

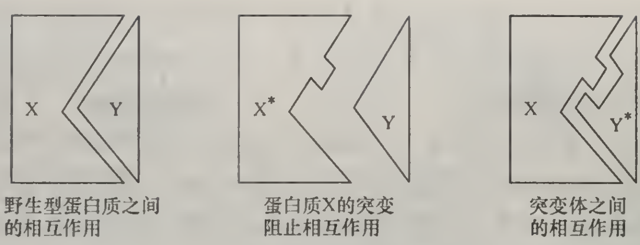


图 21.8　基因外抑制子分析

范围的筛选[51]。如果得到一个温度敏感型的突变型,可以把它转移到限制温度下选择抑制基因。目前已被用来研究线虫运动、产卵、性别决定等过程。

此方法的局限性在于它只能应用在简单生物上,如噬菌体、细菌、酵母、线虫、果蝇等。除了所要研究的目的基因外,还需要一些与研究相关的有价值的突变体。例如,在研究蛋白质相互作用时,如果原突变没有影响到相互作用的结构域,那么抑制基因的分析是无意义的。

2. 合成致死效应

单个基因的突变不能产生致死效果,而两个基因同时突变,则会产生合成致死效果,说明这两个基因所编码的蛋白质在功能上可能是相关的(图 21.9)。Dobzhansky 首次在果蝇中报道了合成致死效应[52],但用其研究蛋白质相互作用在酵母中却应用得最成功。其中一个方法叫做"菌落-分区法"[53]。在此方法中,包含某质粒的酵母细胞是红色的,而不含该质粒的细胞是白色的。如果质粒的存在对于酵母的生长不是必需的,那么菌落呈现红白两种颜色;若为必需的,则只有红色。Bender 和 Pringle 利用这种方法鉴别出与 *MSB1* 基因(与芽形成有关)相关的两个新基因 *BEM1* 和 *BEM2*(与细胞极性、出芽有关)[54]。

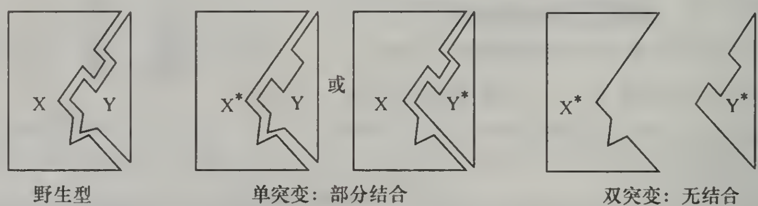


图 21.9　合成致死效应

该方法的优点在于它是一个体内实验,可以在全基因组范围内进行筛选[55]。缺点是工作量大,而且只适于低等生物。

3. 从相关的 mRNA 表达来推测

在不同的生理条件下检测 mRNA 水平,如果一些基因对于特定的条件有相同的转录活性,那么这些基因可归为一类。能够发生相互作用的蛋白质的基因在同一基因类群中富集。Ge 等人选取了酵母中 335 对相互作用的蛋白质,通过计算相互作用密度(protein

interaction density,PID)后发现,同一类群中的 PID 比不同类群之间的 PID 高 30%左右[56]。

该方法的优点是:①将转录组学和相互作用组学联系在一起;②由于 mRNA 的水平是在不同生理条件下测定的,所包含的条件范围比其他技术更广。但这是一个间接的方法。另外,由于涉及到聚类算法(cluster),因此实验结果依赖于不同的算法和参数的选择。

(四) 基于质谱的方法

1. 串联亲和纯化(tandem affinity purification,TAP)[57,58]

此方法的巧妙之处在于设计了一个双重的分子标签,包括 A 蛋白、TEV 蛋白酶切割位点、钙调素结合蛋白(图 21.10)。将 TAP 标签构建到靶蛋白上,然后在宿主细胞内表达融合蛋白,表达水平接近于靶蛋白的天然状态。可以与靶蛋白发生相互作用的其他蛋白质结合到融合蛋白上,形成复合体。细胞裂解物和 IgG 基质温育在一起,通过 A 蛋白复合体结合到 IgG 上。冲洗后,加入 TEV 蛋白酶,洗脱下复合体。在 Ca^{2+} 存在的情况下,洗脱液与包被着钙调素的珠子温育在一起,复合体就结合到珠子上。进一步洗去杂质后,加入 EGTA 螯合,复合体脱落。SDS-PAGE 分离复合体各组分,胰酶酶切后,进行质谱分析。

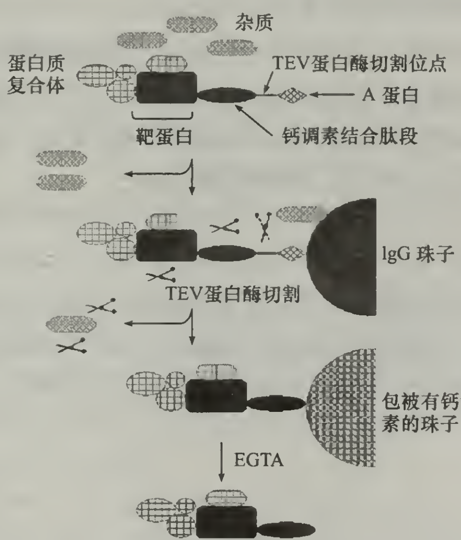


图 21.10 TAP 示意图

2. 高通量质谱蛋白质鉴定(high-throughput mass-spectrometric protein complex identification HMS-PCI)[59,60]

使用一分子标签如 FLAG 标记靶蛋白,使融合蛋白在细胞内过量表达(图 21.11)。通过免疫亲和(FLAG 抗体)捕获到抽提物中靶蛋白形成的蛋白复合体。SDS-PAGE 分离复合体各组分,胰酶酶切后,进行质谱分析。

HMS-PCI 与 TAP 的最大的不同之处是融合蛋白在细胞内的表达水平不同。如果天然状态下靶蛋白在细胞内的浓度过低不足以分析时,可采用 HMS-PCI。但是当靶蛋白过

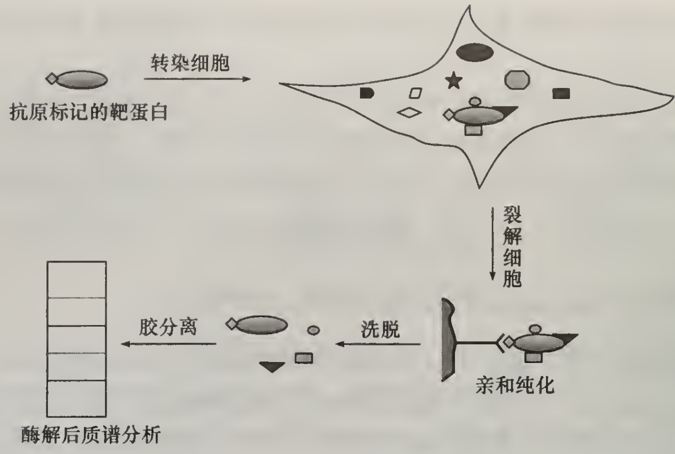


图 21.11　HMS-PCI 示意图

表达时,又会产生假阳性的结果。因此,不同的实验对象应选用合适的实验方法。

(五) 通过基因组分析的软件预测

1. 系统进化谱(phylogenetic profiling)

它主要基于以下两点证据:①在原核生物中,相互作用的蛋白质常常由保守的操纵子所编码[61~63];②相互作用的蛋白质有相似的进化图,即它们在基因组中同时出现或不出现[64,65]。由于使用预测软件,因此该方法的优点是方便、快速。缺点是:①仅适用于原核生物;②在分析之前,需要已知蛋白质之间的同源关系。

2. 染色体邻近

此方法的基础是如果两基因之间的距离小,那么它们在不同的基因组中的位置是保守的。选择的倾向性使一些基因的相邻性得到保留。在一个基因群中,基因是保留的,但其位置则不一定,进化选择的压力倾向使相关功能的基因成簇。由此可以推测蛋白质的功能和相互作用。

3. 基因融合

物种在进化过程中,选择压力使得一些基因融合在一起,它们能够同时被调节,而这些基因所编码的蛋白质在功能上往往是相互联系的[63,66]。Anton 等人在 *Escherichia coli*、*Haemophilus influenzae*、*Methanococcus jannaschii* 全基因组序列的基础上,利用此方法建立了蛋白质相互作用图谱[67]。

(六) 基于蛋白质分子空间结构的预测

1. InterPreTS(protein interaction prediction through tertiary structure)**软件预测**

Aloy 等人建立了一个蛋白质相互作用结构域的数据库(a database of interacting domains, DBID)[68],包含蛋白质复合体三维构象的信息。在进行软件预测时,先搜索 DBID

中与待预测的一对序列相同源的序列，如果存在同源序列，那么根据待测序列在相互作用位点原子的保守性进行打分。该软件网址为：http://www.russell.embl.de/interprets[69]。

2. 补丁分析(patch analysis)

蛋白质相互作用的位点具有特定的性质[70,71]，通过分析蛋白质结构表面的残基，可预测蛋白质相互作用的位点。每个残基需要分析溶解性、疏水性、极性等6个参数。通过6个参数的计算，给出蛋白质相互作用的可能性[72]。

三、结　　语

在研究蛋白质相互作用时，不同的方法有各自的优缺点，采用哪种方法来研究应根据具体的研究对象而定。例如，对信号通路的研究，涉及到短暂的相互作用，可以用酵母双杂交系统；而对于蛋白质复合体的研究则可以采用质谱技术。Mering 等人[73]列举了酵母双杂交、质谱等5种方法研究酵母细胞中蛋白质相互作用的结果，共有80 000对相互作用，但大约只有2 400对相互作用被两种或两种以上的技术所支持。由于方法本身的限制性与分析条件的特异性，往往会造成假阳性或假阴性的结果，因此，需要采用其他不同方法来验证单一方法研究的结果。

第二节　蛋白质芯片

长期以来，人们用来研究蛋白质的方法大多是一次只研究一个或较少的几个相互关联蛋白质，这就导致了要花费几年或者更长的时间去纯化一个蛋白质和确定一个蛋白质的物理化学性质及其生物学功能，这种慢速的、单一的蛋白质功能研究显然不能适应在大科学时期对蛋白质科学研究高效、高通量的要求，而且这种较孤立的研究也不利于我们对生物这个有机整体进行系统的理解。随着基因组计划的实施和逐步完成，对多种相关蛋白质和组织器官蛋白质进行研究已成为可能，而功能蛋白质组研究又对高效、高通量的蛋白质功能研究产生了更加迫切的要求。

生物体生理功能的调节归根结底是由蛋白质来完成的，而传统的蛋白质分析方法慢速、低效并不能适应当前的需要，基因芯片方法虽然可以对基因功能进行高效分析，但是它却是间接对蛋白质进行分析，由于 mRNA 的表达水平和蛋白质的表达水平有很大的差异，因此基因芯片也并不能很好地体现蛋白质和机体生理功能之间的关系。蛋白质芯片技术的出现及发展，并以其微型化、高效、高通量和对蛋白质分析更加直接的特点(表21.1)，成为当今蛋白质研究的热门技术之一。目前已经被广泛应用在疾病诊断，药物设计，研究蛋白质-蛋白质、蛋白质-核酸、蛋白质-脂质体等相互作用，特异性蛋白质的筛选和功能蛋白质组研究等多个领域，并且以自身的特点在更广泛的领域内展示了其重要的作用。在本节内容中将对蛋白质芯片的分类及其特点、检测手段和这种技术的应用作一介绍。

表 21.1　几种研究方法的比较

类　别	通　量	效　率	研究方式	微型化
蛋白质芯片	高通量	高	直接对蛋白质进行分析	是
基因芯片	高通量	高,但不准确	间接对蛋白质进行分析	是
电泳技术	低通量	低	直接对蛋白质进行分析	否
色谱技术	较高通量	较高	直接对蛋白质进行分析	否
质谱技术	高通量	高	直接对蛋白质进行分析	否

一、蛋白质芯片的分类[74~76,84]

迄今为止已经有几种蛋白质芯片被设计使用了,包括在玻璃表面(glass slide)构建的蛋白质芯片,在多孔凝胶垫上(porous gel pad slide)构建的蛋白质芯片和在微孔板上(microwell)构建的蛋白质芯片[74~76,84](图 21.12)。在此分述如下。

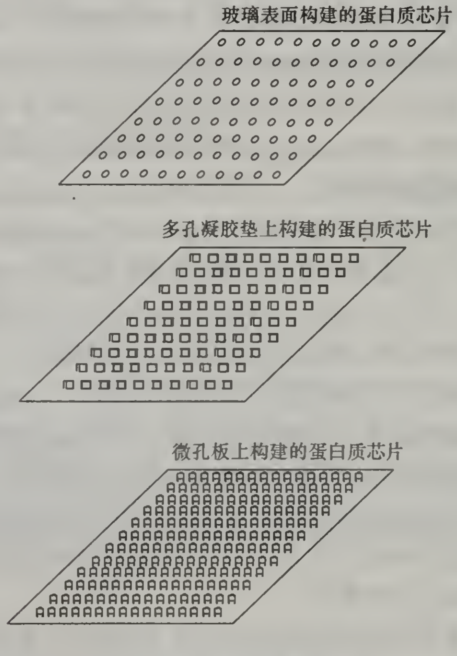


图 21.12　在不同材料上构建的蛋白质芯片

(一) 在玻璃表面构建的蛋白质芯片

在玻璃表面构建蛋白质芯片的一般方法是:首先在玻璃表面上铺以适当的基质;再将蛋白质或者抗体用自动点样仪或者人工的方法分别点在基质上;然后通过一定的条件,使蛋白质样品或者抗体能够锚定在玻璃板上。

在玻璃表面铺上适当的基质后,使平面形成许多有序排列的管、腔或相对独立的点,蛋白质芯片应具有一个亲水的表面。现在经常使用的覆盖玻璃表面的基质有多聚左旋赖氨酸、聚丙烯酰胺等,基质应该使蛋白质可以很好地固定在玻璃表面而且要求噪声较低,但是现在所应用的基质并不能完全达到要求。玻璃表面构建的蛋白质芯片可以制备标准

的微距阵,并可以用基因芯片扫描仪来扫描。因此,适用于自动化、大通量的在线分析,但是这种芯片容易发生交叉污染,在实验过程中应该尽量小心。

MacBeath[77]应用这种芯片分别检测了已知的 3 个蛋白质-蛋白质相互作用、3 个蛋白质-配体作用和 3 个酶-底物的反应,取得了理想的结果。说明了这种蛋白质芯片技术的可行性。他们首先用含有乙醛的硅烷化反应物处理玻璃表面,使之能够和蛋白质的一级胺形成西佛碱连接。蛋白质样品混在含有 40%甘油的磷酸盐缓冲液中刻在玻璃板上,由于甘油的作用,蛋白质保持水合形式可以防止蛋白质脱水变性。这样处理 3h 后,玻璃片用牛血清蛋白(BSA)处理,这不仅可以去除没有反应的乙醛,也可以减少以后步骤中蛋白质之间的非特异性连接。但是由于 BSA 掩盖了一些相对分子质量较小的蛋白质和肽的信号,这种芯片对于相对分子质量较小的蛋白质或者肽的检测能力并不是很高。在实验中发现,可以用 BSA-氮羟基琥珀酰亚胺(BSA-NHS)处理的玻璃片来检测相对分子质量较小的蛋白质和肽。与 BSA 相比,BSA-NHS 单分子性能更好,可以减少蛋白质彼此间的污染。

Ward 等[78]也利用玻璃表面构建的蛋白质芯片来检测几个蛋白质之间的相互作用。他们将抗体与一段寡聚核苷酸相连接,用 RCA(the rolling circle DNA amplification)反应来检测蛋白质之间的相互作用。RCA 反应的原理是:将抗体与一段寡聚核苷酸链相连接后使之与待检测物作用,然后洗脱游离的和非特异性连接的抗体。再加入 DNA 聚合酶、核苷酸和环状 DNA,使环状 DNA 与抗体上的寡聚核苷酸链相连接并在聚合酶的作用下使寡聚核苷酸链延长,当反应进行一段时间后,加入荧光或放射性的标记物来检测。由于核苷酸链的延长可以使信号得到很大的放大,因此可以达到很高的灵敏度。

除上述的几种用来作为玻璃表面基质的材料外,也可以利用多聚赖氨酸[79]、环氧树脂[86]、聚偏乙烯二氟酯(PVDF)[81,82]等作为玻璃表面蛋白质芯片的基质,它们的原理也相似,都是利用聚合反应相互结合成为滤膜用来固定蛋白质或者抗体。在玻璃表面构建蛋白质芯片的方法还有金膜表面[80]光刻板印刷法,它可以与蛋白质高强度结合,而只有很低的背景,并且可以与质谱或者 SPR(surface plasmon resonance)联用。这些检测方法的特点见表 21.2。

表 21.2 蛋白质芯片部分表面材料的比较

表面材料	联结方式	优 点	缺 点
多聚赖氨酸	吸附作用	不需要蛋白质的修饰	非特异性吸附
乙醛	共价连接	高密度、高强度的蛋白质连接能力,高分辨率	表面蛋白质的连接是随机的,容易掩盖蛋白质本身检测位点
环氧树脂	共价连接	蛋白质连接紧密并且具有高分辨率	表面蛋白质的连接是随机的,容易掩盖蛋白质本身检测位点
金膜表面	共价连接	高密度、高强度的蛋白质连接能力,低背景,并可以与质谱、SPR 等配合使用	表面蛋白质的连接是随机的,容易掩盖蛋白质本身检测位点,制作困难,不利于商业化生产
多聚偏乙烯二氟酯	吸附作用	不需要蛋白质的特异性修饰,结合蛋白质能力强	表面蛋白质的连接是随机的,容易掩盖蛋白质本身检测位点

（二）在多孔凝胶垫上构建的蛋白质芯片

与在玻璃表面构建的二维的蛋白质芯片相比，在多孔凝胶垫上构建的蛋白质芯片是一个三维的芯片，它比二维的芯片有更大的容积，可达二维芯片的 100 倍左右。而且还可以更好地固定蛋白质，减少蛋白之间的交叉污染。

Mirzabekov[83]在其实验中将 3%的丙烯酰胺与交联剂 N,N'-(1,2-二羟亚乙基)双丙烯酰胺(DHEBA)的混合物铺在硅烷处理过的玻片表面。加入 DHEBA 可以提高凝胶的多孔性。处理后通过有许多 100 μm×100 μm 的方形小罩，把聚丙烯酰胺暴露在紫外线下，加速凝胶的聚合。然后将未聚合的分子用水冲洗去除。蛋白质或者抗体可以通过戊二醛激活凝胶或者通过酰肼基部分代替凝胶中的酰胺基将蛋白质或抗体固定在凝胶中。Mirzabekov[83]用此方法检测了一些酶-底物反应和抗原-抗体的相互作用，验证了其可行性。

由于凝胶支持物的应用，这种蛋白质芯片具备了新的特点。首先它具有更大的容积，而且三维蛋白质芯片更好固定蛋白的性质可有效地减少蛋白质的交叉污染，在凝胶的均一水环境中可以极大地减少由于脱水而造成的蛋白质变性。但是这种芯片却不易改变它的缓冲条件去适应不同的检测物质，而且在玻片的处理上也比二维蛋白质芯片要复杂得多，价格也更高。

（三）在微孔板上构建的蛋白质芯片

Mendoza 等[85]在传统微滴定板的基础上加以改进，利用机械手在 96 孔的每一个孔的底部点上同样的 4 组蛋白质，每组 36 个点，含有 8 种不同抗原和标记的蛋白质。然后用配套的免疫分析仪分析检测，再通过 CCD 扫描监测，制成了在微孔板上构建的蛋白质芯片并验证了其可行性。

Zhu 等[86]发展了这种蛋白质芯片技术对酵母的 119 个蛋白激酶进行了高通量的生物化学活性分析。他们先用多聚二甲基硅氧烷作为蛋白质芯片表面的支持剂，将 17 种不同的底物点在各自在蛋白质芯片的专门的位点上，然后分别用 119 种蛋白激酶和放射性标记的 ATP 混合物与其一起反应；反应一段时间后，洗去蛋白激酶和没有特异性连接的 ATP，最后将蛋白质芯片检测的结果进行分析。在实验中产生了意想不到的结果，他们发现在酵母中有 1/4 的磷酸化激酶可以磷酸化酪氨酸，并且丝氨酸-苏氨酸激酶家族的成员也不例外。他们的工作也证实了蛋白质芯片技术在高通量蛋白质分析中应用的可行性。

在微孔板上构建的蛋白质芯片也有其特点。在微孔板中蛋白质分离比较好，不容易发生交叉污染，微孔板中均一的水环境也使得蛋白质不易脱水变性，这对蛋白质芯片的检测准确性是非常重要的，而且这种蛋白质芯片检测灵敏度也是很高的。在微孔板上构建的蛋白质芯片可以在反应的过程中改变不同的缓冲条件去适应反应的需要，在个别的小孔中加入特别的物质都不会对检测的结果产生影响，而且价格也很便宜。

在玻璃表面构建的蛋白质芯片、在多孔凝胶垫上构建的蛋白质芯片和在微孔板上构建的蛋白质芯片都已广泛应用于科学研究和大规模的生产中。表 21.3 将这三种蛋白质芯片的特点进行了比较。

表 21.3　不同种类的蛋白质芯片的特点

类　别	特　点
玻璃表面构建的蛋白质芯片	可以与基因芯片的制作和检测工具配套使用；很容易蒸发；不适合用于多步反应；造价便宜；容量较小；容易发生交叉污染
多孔凝胶垫上构建的蛋白质芯片	可以与基因芯片的制作和检测工具配套使用；不易蒸发；可以用于多步反应，但是有其较固定的缓冲条件；造价昂贵；容量较大；不易发生交叉污染
在微孔板上构建的蛋白质芯片	可以与基因芯片的制作和检测工具配套使用；不易蒸发；适用于多步反应；造价便宜；容量较大；不易发生交叉污染

二、蛋白质芯片的检测方法

蛋白质芯片的种类繁多，其检测手段更是层出不穷，经过多年的发展可分为以下三类：探针标记检测法、无探针标记检测法和动态分析检测法。

（一）探针标记检测法

探针标记检测法可以根据标记探针的特点大致分为两种：荧光标记的探针和放射性同位素标记的探针。

荧光检测方法是目前最常用的检测方法，它具有简单、安全和高灵敏度的特点[87,88]。荧光检测法也可以和基因芯片的扫描仪配合使用，这样可以提高分析效率，适于大量数据的采集和分析。

现在常用的荧光标记方法是一步标记法和二步标记法。一步标记法就是用荧光蛋白或者可以发光的小分子经过一步将蛋白质作标记检测。而二步标记法就是先用一个可检测到的标签探针将蛋白质标记后，再在第二步操作中用一个合适的发光试剂作标记。例如，先用生物素标记，然后用抗生物素蛋白与之相互作用来检测，由于通过标记后再检测可以起到信号放大的作用，所以，这种方法有较高的灵敏度。但是二步标记法使操作步骤变得复杂，而且有的蛋白质也不适合此种方法。

还有一种荧光标记的方法是 RCA 检测法[78]，由于此种方法可以造成信号的放大，因此具有较高的灵敏度。

放射性同位素探针标记检测法与荧光探针标记检测法原理相似，这种方法目前应用很广泛。Zhu 等[88]运用此方法分别检测了蛋白质-蛋白质、蛋白质- DNA、蛋白质-药物和酶-底物的相互作用。

（二）无探针标记的检测方法

由于探针和蛋白质的相互作用可能造成某些蛋白质的空间结构的变化，影响蛋白质的活性，从而导致检测结果的准确性下降。为此，有人发展了一些非探针标记的检测方法。运用这些方法可以直接对蛋白质芯片进行检测，在实验条件的改进方面这对于一些蛋白质的检测是非常重要的，但是这些方法仍然处于发展中。

有人已经运用 SELDI（surface-enhanced laser desorption/ionization）质谱来检测低密度

蛋白质芯片[89]。SELDI 是在 MALDI(matrix-assisted laser desorption/ionization)基础上改进后实行表面增强的飞行质谱。在一个金属表面的蛋白质芯片上,结合的蛋白质用激光激发后进入质谱进行数据分析。对 SELDI 来说,可以直接将血清、尿液、组织抽取物等不需处理直接点样检测;由于一部分非特异结合的分析物被洗去,因而出现的质峰非常一致,有利于后期分析。与二维电泳相比,它可以检测到低丰度的蛋白质,所需要时间短,而且有利于自动化控制。

原子力显微镜(atomic force microscopy, AFM)也可以作为蛋白质芯片的检测工具[80]。原子力显微镜具有一个对力非常敏感的微悬臂,其尖端有一个微小的探针,当探针轻微地接触样品表面时,由于探针尖端的原子与样品表面的原子之间产生极其微弱的相互作用力而使微悬臂弯曲,将微悬臂弯曲的形变信号转换成光电信号并进行放大,就可以得到原子之间力的微弱变化的信号。从这里我们可以看出,原子力显微镜设计的高明之处在于利用微悬臂间接地感受和放大原子之间的作用力,从而达到检测的目的。原子力显微镜同样具有原子级的分辨率,而且原子力显微镜既可以观察导体,也可以观察非导体。根据这个原理,可以运用原子力显微镜检测蛋白质芯片结合前后表面拓扑学性质的变化来检测蛋白质芯片。

(三) 动态的分析检测方法

有时为了研究抗原-抗体相互作用的动力学变化,需要一种可以实施的即时的检测方法。表面激源共振生物传感器(surface plasmon resonance, SPR)检测法的原理是:利用表面等离子共振现象监测传感片表面介质的折射率变化,而这一变化仅与传感片表面结合的生物大分子的量成正比,溶液中的分子不会干扰,因此非常特异、敏感。现在已经成为一个比较成熟的多方面研究受体-配基相互作用动力学的检测工具[90~92]。Sapsford 等[93]发展了一个抗体芯片的生物传感器去研究抗体作用的动力学。这种检测方法可在直径只有 200im 的点上检测到重要的信号。因此,可以期待在不久的将来可采用这种方法做大规模的平行的动力学分析。

三、蛋白质芯片技术的应用及发展前景

由于蛋白质芯片技术其自身的特点,它已经在广泛的领域中得到了应用。在疾病诊断方面,它比传统的检测方法需要更短的时间而只需要更少的样品量,但是却能给出更多的诊断信息,而且它还可以用来进行自我检测,因此可以成为以后医疗诊断的发展方向。而在药物筛选方面,蛋白质芯片技术可以用来寻找药物的靶标,检查药物的毒副作用,而且用蛋白质芯片做大规模的筛选还可以省略大量的动物试验,从而缩短药物筛选所需要的时间。而在研究蛋白质相互作用、靶蛋白的筛选方面,蛋白质芯片更是以其直接对蛋白质进行分析的优势,发挥着极大的作用。随着人类基因组计划的逐步完成,蛋白质组计划的实施,对生物体某组织、器官、甚至全生物体组蛋白质组的研究已经成为了发展方向,而在其中蛋白质芯片更能发挥其优势。

当然,蛋白质芯片技术作为一项正在发展中的技术,其本身还存在很多问题。例如,作为芯片表面基质的材料的发展,作为蛋白质芯片的特异性抗体的筛选制作,检测方法的

改进等，这些都是当今困扰蛋白质芯片发展的问题，也是需要不断解决的问题。

（唐柳娅　马丹军）

参考文献

1 Trevor CD. Signal transduction by the Wnt family of ligands. Biochem J, 1998, 329:209～223

2 Smith DB, Johnson KS. Single-step purification of poly peptides expressed in *Escherichia coli* as fusions with glutathione S-transferase. Gene, 1988, 67:31～40

3 Mayer BJ, Jackson PK, Baltimore D. The noncatalytic src homology region 2 segment of *abl* tyrosine kinase binds to tyrosine-phosphorylated cellular proteins with high affinity. Proc Natl Acad Sci USA, 1991, 88:627～631

4 Weng Z, Taylor JA, Turner CE *et al*. Detection of Src homology 3-binding proteins, including paxillin, in normal and v-Src-transformed Balb/c 3T3 cells. J Bio Chem, 1993, 268:14956～14963

5 Zhang XF, Settleman J, Kyriakis JM *et al*. Normal and oncogenic $p21^{ras}$ proteins bind to the amino-terminal regulatory domain of c-Raf-1. Nature, 1993, 364:308～313

6 Flynn DC, Leu TH, Reynolds AB *et al*. Identification and sequence analysis of cDNAs encoding a 110-kilodalton actin filament-associated $pp60^{src}$ substrate. Mol Cell Biol, 1993, 13:7892～7900

7 Hu P, Margolis B, Skolnik EY *et al*. Interaction of phosphatidylinositol 3-kinase-associated p85 with epidermal growth factor and platelet-derived growth factor receptors. Mol Cell Biol, 1992, 12:981～990

8 Formosa T, Barry J, Alberts BM *et al*. Using protein affinity chromatography to probe structure of protein machines. Methods Enzymol, 1991, 208:24～25

9 Lazarides E, Lindberg U. Actin is the naturally occurring inhibitor of deoxyribonuclease I. Proc Natl Acad Sci USA, 1974, 71:4742～4746

10 Soutar AK, Wade DP. In: Creighton, T ed. Protein function: a practical approach. Ligand bloting, Oxford: IRL Press, 55～76

11 Whyte P, Williamson NW, Harlow E. Cellular targets for transformation by the adenovirus E1A proteins. Cell, 1989, 56:67～75

12 Ewen ME, Faha B, Harlow E *et al*. Interaction of p107 with cyclin A independent of complex formation with viral oncoprotein. Science, 1992, 255:85～87

13 Faha B, Ewen ME, Tsai LH *et al*. Interaction between human cyclin A and adenovirus E1A-associated p107 protein. Science, 1992, 255:87～90

14 Harlow E, Whyte P, Franza BR *et al*. Association of adenovirus early-region 1A proteins with cellular polypeptides. Mol Cell Biol, 1986, 6:1579～1589

15 Denny JB, Blobel B. ^{125}I-labeled crosslinking reagent that is hydrophilic, photoactivatable, and cleavable through an azo linkage. Proc Natl Acad Sci USA, 1984, 81:5286～5290

16 Chen Y, Ebright YW, Ebright RH. Identification of the target of a transcription activator protein by protein-protein photocrosslinking. Science, 1994, 265:90～92

17 Bleil JD, Wassarman PM. Identification of a ZP3-binding protein on acrosome-intact mouse sperm by photoaffinity crosslinking. Proc Natl Acad Sci USA, 1990, 87:5563～5567

18 de Gunzberg J, Richl R, Weinberg RA. Identification of a protein associated with p21ras by chemical crosslinking. Proc Natl Acad Sci USA, 1989, 86:4007～4011

19 Alexander S, Maria MC, Huang FT *et al*. Interaction of EGF receptor and Grb2 in living cells visualized by fluorescence resonance energy transfer (FRET) microscopy. Curr Biol, 2000, 10:1395～1398

20 Li HY, Ng EKO, Lee SMY *et al*. Protein-protein interaction of FHL3 with FHL2 and visualization of their interaction by green fluorescent protein (GFP) two-fusion fluorescence resonance energy transfer (FRET). J Cell Biochem, 2001, 80:293～303

21 Miyawaki A, Llopis J, Heim R *et al*. Fluorescent indicators for Ca^{2+} based on green fluorescent proteins and calmodulin. Nature, 1997, 388:882~887

22 Fivash M, Towler EM, Fisher RJ. BIAcore for macromolecular interaction. Curr Opin Biotech, 1998, 9:97~101

23 Krone JR, Nelson RW, Dogruel D *et al*. BIA/MS: Interfacing biomolecular interaction analysis with mass spectrometry. Anal Biochem, 1996,244:124~132

24 Tamamura H, Ishihara T, Otaka A *et al*. Analysis of the interaction of an anti-HIV peptide, T22 ([Tyr5, 12, Lys7]-polyphemusin II), with gp120 and CD4 by surface plasmon resonance. Biochim Biophys Acta, 1996,1298:37~44

25 Schier R, Bye J, Apell G *et al*. Isolation of high-affinity monomerichuman anti-c-erbB-2 single chain Fv using affinity-drivenselection. J Mol Biol, 1996,255:28~43

26 Medina MB. Hygromycin B antibody production and characterization by a surface plasmon resonance biosensor. J Agricultural Food Chem, 1997, 45:389~394

27 Oddie GW, Gruen LC, Odgers GA *et al*. Identification and minimization of nonideal binding effects in BIAcore analysis: ferritin/anti-ferritin Fab' interaction as a model system. Anal Biochem, 1997, 244:301~311

28 Myszka DG, Morton TA, Doyle ML *et al*. Kinetic analysis of protein antigen-antibody interaction limited by mass transport on an optical biosensor. Biophys Chem, 1997, 64:127~137

29 Autiero M, Gaubin M, Mani JC *et al*. Surface plasmon resonance analysis of gpl 7, a natural CD4 ligand from human seminal plasma inhibiting human immunodeficiency virus type-1 gp120-mediated syncytium formation. Eur J Biochem, 1997,245:208~213

30 Rock FL, Peterson D, Weig BC *et al*. Binding of leptin to the soluble ectodomain of recombinant leptin receptor: a kinetic analysis by surface plasmon resonance. Horm Metab Res, 1996, 28:748~750

31 Alam SM, Travers PJ, Wung JL *et al*. T-cell receptor affinity and thymocyte positive selection. Nature, 1996, 381: 616~620

32 Chen BJ, Hage 1", Sebald W. Global and local determinants for the kinetics of interleukin-4/interleukin-4 receptor á-chain interaction: a biosensor study employing recombinant interleukin-4 binding protein. Eur J Biochem, 1996, 240: 252~261

33 Wu H, Myszka DG, Tendian SW *et al*. Kinetic and structural analysis of mutant CD4 receptors that are defective in HIM gp120 binding. Proc Natl Acad Sci USA, 1996, 93:15030~15035

34 Cheskis B, Freedman LP. Modulation of nuclear receptor interactions by ligands: kinetic analysis using surface plasmon resonance. Biochemistry, 1996, 35:3309~3318

35 Jensen KK, Orum H, Nielsen PE *et al*. Kinetics for hybridization of peptide nucleic acids (PNA) with DNA and RNA studied with the BIAcore technique. Biochemistry, 1997, 36:5072~5077

36 MacKenzie CR, Hirama T, Lee KK *et al*. Quantitative analysis of bacterial toxin affinity and specificity for glycolipid receptors by surface plasmon resonance. J Biol Chem, 1997, 272:5533~5538

37 Lackmann M, Bucci T, Mann PJ *et al*. Purification of a ligand for the EPH-like receptor HEK using a biosensor- based affinity detection approach. Proc Natl Acad Sci USA, 1996, 93:2523~2527

38 Fields S, Song OK. A novel genetic system to detect protein-protein interactions. Nature, 1989, 340:245~246

39 Boeke JD, Walkwan C, Sung PG. A positive selection for mutants lacking orotidine-5'-phosphate decarboxylase activity in yeast: 5-fluoro-orotic acid resistance. Mol Gen Genet, 1984,197(2):345~346

40 Vidal M, Brachmann RK, Fattaey A *et al*. Reverse two-hybrid and one-hybrid systems to detect dissociation of protein-protein and DNA-protein interactions. Proc Natl Acad Sci USA, 1996, 93:10315~10320

41 Shih HM, Goldman PS, DeMaggio AJ *et al*. A positive genetic selection for disrupting protein-protein interactions: Identification of CREB mutations that prevent association with the coactivator CBP. Proc Natl Acad Sci USA, 1996, 93:13896~13901

42 White MA. The yeast two-hybrid system: forward and reverse. Proc Natl Acad Sci USA, 1996, 93:1001~1003

43 Tirode F, Malaguti C, Romero F *et al*. A conditionally expressed third partner stabilized or prevents the formation of a transcriptional activator in a three-hybrid system. J Biol Chem, 1997, 272:22995~22999

44 Printen JA, Sprague G. Protein-protein interactions in the yeast pheromone response pathway: Ste5p interacts with all members of the MAP kinase cascade. Genetics, 1994, 138:609~619

45 Zhang J, Lautar S. A yeast three-hybrid method to clone ternary protein complex components. Anal Biochem, 1996, 242:68~72

46 Tomashek JJ, Sonnenburg JL, Artimorich JM *et al*. Resolution of subunit interactions and cytoplasmic subcomplexes of the yeast vacuolar protein-translocating ATPase. J Biol Chem, 1996, 271:10397~10404

47 Ma J, Ptashne M. A new class of yeast transcriptional activators. Cell, 1987, 51:113~119

48 Smith GP. Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. Science, 1985, 228:1315~1317

49 Il'Ichev AA, Minenkova OO, Tatkov SI *et al*. M13 filamentous bacteriophage in protein engineering. Dokl. Akad. Nauk SSSR, 1989, 307:431~433

50 Devlin JJ, Panganiban LC, Devlin PE. Random peptide libraries: a source of specific protein binding molecules. Science, 1990, 249:404~406

51 Hodgkin J, Kondo K, Waterson RH. Suppression in the nematode *Caenorhabditis elegans*. Trends Genet, 1987, 3:325~329

52 Dobzhansky T. Genetics of natural populations. XIII. Recombination and variability in populations of *Drosophila pseudoobscura*. Genetics, 1946, 31:269~290

53 Hieter P, Mann C, Snyder M *et al*. Mitotic stability of yeast chromosomes: a colony color assay that measures nondisjunction and chromosome loss. Cell, 1985, 40:381~392

54 Bender T, Pringle JR. Use of a screen for synthetic lethal and multicopy suppressee mutants to identify two new genes involved in morphogenesis in *Saccharomyces cerevisiae*. Mol Cell Biol, 1991, 11:1295~1305

55 Tong AH, Evangelista M, Parsons AB *et al*. Systematic genetic analysis with ordered arrays of yeast deletion mutants. Science, 2001, 294:2364~2368

56 Ge H, Liu Z, Church GM *et al*. Correlation between transcriptome and interactome mapping data from *Saccharomyces cerevisiae*. Nat Genet, 2001, 29:482~486

57 Rigaut G, Shevchenko A, Rutz B *et al*. A generic protein purification method for protein complex characterization and proteome exploration. Nat Biotechnol, 1999, 17:1030~1032

58 Gavin AC, Bche M, Krause R *et al*. Functional organization of the yeast proteome by systematic analysisi of protein complexes. Nature, 2002, 415:141~147

59 Ho Y, Gruhler A, Heilbut A *et al*. Systematic identification of protein complexes in *Saccharomyces cerevisiae* by mass spectrometry. Nature, 2002, 415:180~183

60 Mann M, Hendrickson RC, Pander A. Analysis of proteins and proteomes by mass spectrometry. Annu Rev Biochem, 2001, 70:437~473

61 Huynen M, Snel B, Lathe W3rd *et al*. Predicting protein function by genomic context: quantitive evaluation and qualitative inferences. Genome Res, 2000, 10:1204~1210

62 Overbeek R, Fonstein M, D'Souza M *et al*. The use of gene clusters to infer functional coupling. Proc Natl Acad Sci USA, 2001, 96:2896~2901

63 Dandekar T, Snel B, Huynen M *et al*. Conservation of gene order: a fingerprint of proteins that physically interact. Trends Biochem Sci, 1998, 23:324~328

64 Pellegrini M, Marotte EM, Thompson MI *et al*. Assigning protein functions by comparative genome analysis: protein phylogenetic profiles. Proc Natl Acad Sci USA, 1999, 96:4285~4288

65 Huynen MA, Bork P. Measuring genome evolution. Proc Natl Acad Sci USA, 1998, 95:5849~5856

66 Tamames J, Casari G, Ouzounis C *et al*. Conserved clusters of functionally related genes in two bacterial genomes. J Mol Evol, 1997, 44:66~73

67 Anton JE, Loannis I, Nikos CK *et al*. Protein interaction maps for complete genomes based on gene fusion events. Nature, 1999, 402:86~90

68 Aloy P, Russell RB. Interrogating protein interaction networks through structural biology. Proc Natl Acad Sci USA, 2002, 99(9):5896～5901
69 Aloy P, Russell RB. InterPreTS: protein interaction prediction through tertiary structure. Bioinformatics, 2003, 19(1):161～162
70 Jones S, Thornton JM. Protein-protein interactions: a review of protein dimer structures. Prog Biophys Mol Biol, 1995, 63: 31～65
71 Jones S, Thornton JM. The principles of protein-protein interactions. Proc Natl Acad Sci USA, 1996, 93: 13～20
72 Jones S, Thornton JM. Prediction of protein-protein interaction sites using patch analysis. J Mol Biol, 1997, 272(1): 133～143
73 Mering VC, Krause R, Snel B *et al*. Comparative assessment of large-scale data sets of protein-protein interactions. Nature, 2002, 417:399～403
74 MacBeath G.. Protein microarrays and proteomics. Nat Genet. 2002 Dec; 32 Suppl: 625～632
75 Zhu H, Synder M. Protein arrays and microarrays. Curr Opin Chem Biol, 2001, 5(1):40～45
76 Eickhoff H, Konthur Z, Leuking A *et al*. Protein array technology: the tool to bridge genomics and proteomics. Adv Biochem Eng Biotechnol, 2002, 77:103～112
77 MacBeath G, Schreiber SL. Printing proteins as microarrays for high-throughput function determination. Science, 2000, 289:1760～1763
78 Schweitzer B, witshire S, Ward DC *et al*. Immunoassays with rolling circle DNA amplification: a versatile platform for ultrasensitive antigen detection. Proc Natl Acad Sci USA, 2000,97:10113～10119
79 Haab BB, Dunham MJ, Brown PO. Protein microarrays for highly parallel detection and quantitation of specific proteins and antibodies in complex solutions. Genome Biol, 2001, 2(2):R4
80 Jones V W, Kenseth JR, Porter MD *et al*. Microminiaturized immunoassays using atomic force microscopy and compositionally pattered antigen arrays. Anal Chem, 1998.70(7) 1233～1241
81 Cahill DJ. Protein and antibody arrays and their medical applications. J Immunol Methods, 2001, 250:81～91
82 Walter G, Bussow K,Cahill D *et al*. Protein arrays for gene expression and molecular interaction screening. Curr Opin Microb, 2000, 18:393～397
83 Arenkov, kukhtin A, Gemmell A, *et al*. Protein Microchips: use for immunoassay and enzymatic reaction. Anal Biochem, 2000, 278:123～131
84 Zhu H, Snyder M. Protein chip technology. Curr Opin Chem Biol, 2003, 7(1): 55～63
85 Mendoza L G, McQuary, Mongan A *et al*. High-throughput microarray-based enzymelinked immunosorbent(ELISA). Biotechniques, 1999, 27(4): 778～788
86 Zhu H, Klemic JF, Chang S *et al*. Analysis of yeast protein kinases using protein chips. Nat Genet, 2000, 26:283～289
87 Haab BB. Advance in protein microarray technology for protein expression and interaction profiling. Curr Opin Drug Disc Dev, 2001, 4: 116～123
88 Templin MF, Stoll D, Schrenk M, *et al*. Protein microarray technology. Trends Biotechnol, 2002, 20:160～166
89 Davies H, Lomas L, Austen B. Profiling of amyloid beta peptide variants using SELDI Protein Chip arrays. BioTechniques, 1999, 27:1258～1261
90 McDonnell JM. Surface plasmon resonance: towards an recognition. Curr Opin Chem Biol, 2001, 5: 572～577
91 Nieba L, Nieba-Axmann SA, Persson A *et al*. BIACORE analysis of histidine-tagged proteins using a chelating NTA sensor chip. Anal Biochem, 1997, 252:217～228
92 Salamon Z, Brown MF, Tollin G. Pliasmon resonance spectroscopy: probing molecular interactions within membranes. Trends Biochem Sci, 1999, 24:213～219
93 Sapford DG, Liron Z, Shubin YS *et al*. Kinetics of antigen binding to arrays of antibodies in different sized spots. Anal Chem, 2001, 73: 5518～5524

第二十二章　蛋白质组生物信息学

生物信息学在蛋白质组学的研究中将起到特殊的重要作用，因为基因组和蛋白质组研究提供数据的数量之巨大在生物学上是史无前例的，这非同一般的大量数据必须要有高度自动化的处理，包括数据的获取、输入、储存、加工以及数据库之间的联系。输入和输出数据必须非常迅速并有质量控制，有必要和完备的信息管理系统，数据处理需要设计各种特殊软件，以及对这些软件的整合，形成高通量、自动化、可编辑的各类蛋白质组学分析平台。由于蛋白质组比基因组具有更大的复杂性，因而蛋白质组生物信息学研究更具有挑战性。

蛋白质组生物信息学的研究内容主要包括大规模蛋白质组学实验信息的产生，将实验信息处理成具有生物学意义的实验结果，对实验结果的分析和实验结果的发布，以及对上述过程的管理。

一、大规模蛋白质组学实验信息的产生

蛋白质组实验产生的实验信息大多依赖于仪器，如双向凝胶电泳仪、质谱仪等，这些实验数据的产生当然离不开信息处理和采集软件，不过这些软件基本上与仪器绑定工作，在此不再赘述。

二、将实验信息处理成具有生物学意义的实验结果

生物学实验产生的数据本身大多是数字信号（如谱线、图像或数值），如何把这些数字信号转换成具备生物学意义的实验结果（如蛋白质名称、多肽序列、蛋白质结构、蛋白质差异等）是蛋白质组学生物信息研究的核心。目前的实验信息主要来自两方面，一是蛋白质组表达谱及其相关信息，如双向凝胶电泳图谱、蛋白质芯片图谱、蛋白质相互作用等。另一方面则是以质谱技术为中心的蛋白质鉴定数据。前者大多是图像，所涉及的信息处理技术有图像获取、识别、定量及其差异比对等过程。本书的第十五章已对双向凝胶电泳图谱分析做了阐述。在此主要讨论如何将质谱的蛋白质鉴定数据转换为具有生物学意义的实验结果。

蛋白质样品被选定的酶水解，形成多肽，不同质量的多肽被质谱仪检测出来，得到质谱（MS）。肽可以进一步被打碎，并测得碎片质量分布，即串联质谱（MS/MS）。用质谱识别蛋白质的方法有三种：第一种是肽质量指纹图谱结合数据库搜索识别蛋白质。先对蛋白质模拟酶解，得到理论质谱，再将理论质谱与实验质谱进行比较，符合的作为结果。这样的算法有 Mascot、ProFound 等；第二种是用串联质谱数据搜索数据库。对蛋白质模拟酶解，再对肽模拟碰撞，得到理论串联质谱与实验串联质谱进行比较，符合的作为结果。所采用的方法有 Mascot、SEQUEST、MS-Tag 等；第三种是从串联质谱数据直接推测肽序

列，称作 *de novo* 的方法。此法依据串联质谱的物理意义直接推测肽的氨基酸序列，而不是搜索数据库。但由于实验数据的不完整性，只能得到部分序列信息，还要进一步利用这些部分信息搜索数据库，以鉴定整个蛋白质。所用的方法如：图论、动态规划、局部搜索。这种方法严重依赖于质谱的质量，目前未得到广泛应用。

（一）肽质量指纹图谱结合数据库搜索识别蛋白质

目前普遍采用的质谱对蛋白质识别技术被称为肽质量指纹图谱（peptide mass fingerprint，PMF）。蛋白质经过特异性蛋白酶切后，产生的多肽片段通过质谱测定其分子质量，那么就可以利用这些多肽的实验分子质量同蛋白质数据库中的所有蛋白质，在计算机上与经过同样特异性的蛋白酶理论酶切后，所产生的多肽的理论分子质量进行比较，通过模式识别算法，获得一系列最佳匹配的蛋白质，缩小蛋白质识别的范围，从而最终可能获得与实验一致的蛋白质。因此，可以利用计算机，模拟蛋白质酶切以及测定分子质量的过程，设计一套基于质谱技术的蛋白质的识别及其相关功能查询的软件系统。更进一步，可以利用 MS 所获得的一系列多肽分子质量，通过一个功能强大的查询系统，有针对性地对蛋白质数据库或核酸数据库进行操作。

通过这样一个基于多肽指纹图谱鉴定的计算机分析系统（图 22.1），用户可以选定所要查询的蛋白质序列的范围（分类及分子质量），然后按照用户界定的条件（如多肽的分子质量及精确范围，酶的种类，修饰，最多可能漏掉的酶切位点等），然后根据打分矩阵来判

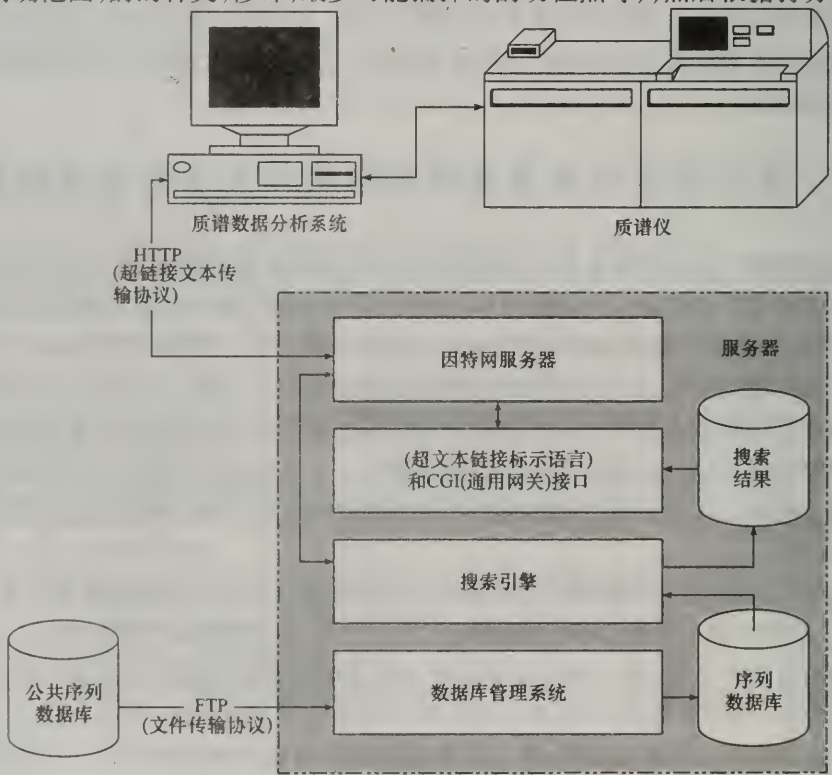


图 22.1　多肽指纹图谱鉴定的计算机分析系统

定哪些符合条件且显著性足够高的蛋白质序列,并按照分值的高低,给用户显示相关的各种信息。下面分别介绍多肽指纹图谱鉴定的计算机分析。

1. 理论酶切数据库结构

(1) 蛋白质数据库的处理

数据库来源包括 Swiss-Prot、TrEMBL 和 TrEMBL-new 所有的非冗余的蛋白质信息。Swiss-Prot 是一个含有注释信息的蛋白质数据库。TrEMBL 则是对 Swiss-Prot 的补充,那些还没有进行整合进入 Swiss-Prot 的翻译的 CDS,都放在 TrEMBL 内,这些 CDS 来源于 EMBL 核酸库。TrEMBL 作为 Swiss-Prot 的预先考虑部分。另外,TrEMBL-new 由 EMBL 核酸库中的 CDS 产生,还未经过整理。TrEMBL-new 要经过注释升级和去除冗余,才能进入 TrEMBL。

(2) 理论酶切多肽数据库的建立

根据每个蛋白酶所具有的特异性识别位点,可以将所有的蛋白质用不同的蛋白酶切后,形成多肽片段。这里要考虑试验中可能出现的不完全酶切的情况。例如,一条蛋白质序列被一种酶完全酶切后,产生 n 个切点,$n+1$ 条片段,那么,如果在最多可能漏掉 3 个酶切位点的情况,会产生 $4n-2$ 条不同的多肽片段。另外一种情况,则是实验中可能会存在用 2 种不同的蛋白酶切同一条蛋白质序列的情况,那么就会出现 2 种酶的各种随机组合的识别位点(当然,某些酶是不可能同时作用于一条蛋白质序列的),产生更多的多肽片段。所有的情况都应考虑进去,将产生的各种不同的多肽片段放入数据库,并与相应的酶、蛋白质序列关联,同时记下产生的条件以及相应的位置等信息,同时计算每一条多肽片段的分子质量。

(3) 分子质量的计算

常用的蛋白质氨基酸一共有 20 种,采用最精确的分子质量来计算多肽的分子质量。由于经常会对蛋白质进行修饰,如糖基化、磷酸化以及酰胺化等,这样蛋白质以及酶切后的多肽的分子质量就会改变。因此,如果当用户选择了某种修饰时,分子质量的计算则会不同。

2. 分子质量查询(MOWSE)打分体系[1]

在建立的多肽数据库中,通常会给每一条多肽分配一个出现频度因子,出现频度因子可如下计算:

计算多肽数据库所有的分子质量在 10 000Da 以内的多肽片段总数 N,再对分布在 10 000Da 框内的每个相隔 100Da 的亚框,计算落在各亚框中的多肽片段总数 n,则落在各亚框中的多肽片段获得相应的出现频度(F)为:

$$F = n/N$$

定义所有亚框中最大出现频度为 F_{max},则得到归一化各亚框中的多肽片段相应的出现频度为:

$$F_n = F/F_{max}$$

利用出现频度因子 F_n,可以增加背景分值,降低随机匹配的影响,有效提高分辨率。所有的匹配片段都将对最终的分值产生贡献,非匹配片段的贡献将被忽略。最终的分值

按下式计算：

$$Score = 50/(P_n * H)$$

P_n: the product of n distribution scores(出现频度因子 F_n)；

H:Hit protein MW in kDa

MOWSE 算法是一个好的起点,因为它正确地将蛋白质被位点特异性的蛋白酶消化的行为模式化。其优点表现在:用简单的规则来判断一个结果是否显著,实现对数据库的快速查询;不同形式的匹配结果可以在同一查询过程中实现;基于不同数据库的不同查询结果可以相互比较;查询参数通过反复实践更容易被优化。利用基于打分算法的概率表明所得结果在统计上的显著性,如果某个确切的且分值较高的匹配,在数据库中并不能找到相符的蛋白质,最高分值的匹配则是与其同源性最高的序列,通常是与这个测试蛋白质相近的物种的同源蛋白质。实验表明,可以通过至少 3～4 条多肽分子质量,扫描 50 000 多条多肽的数据库,来确定样品蛋白质的身份,通过打分算法可以容忍实验中相差几个 Da 的分子质量误差。利用打分方案,可以使分子质量更大的多肽具有更高的得分权重,可以在不同大小的蛋白质中多肽分子质量的非随机分配被补偿,可以减少在较大的蛋白质($>200\times10^3$Da)的随机分值的积累影响。

3. Mascot 系统

Mascot[1]是在 MOWSE 程序的基础上进一步开发的。MOWSE 是第一个使用肽质量指纹(peptide mass fingerprinting)识别蛋白质的程序。MOWSE 主要缺点是:数据库使用索引进行预处理,这些索引往往容量很大并且创建较慢,每次进行数据库更新时操作比较繁琐;不处理可变修饰的情况;提供功能十分有限。与 MOWSE 相比,Mascot 功能大为增强,Mascot 的特点是:①利用多处理器系统处理数据;②完全基于 Web 界面(核心搜索引擎使用 ANSI C 编写,使用 Perl CGI 进行 Web 处理);③去除了对数据库进行索引的处理方法;④可以处理固定和可变修饰的情况;⑤增加了概率打分和支持 MS/MS 数据查询,支持多种流行数据库;⑥数据处理速度大大增加。主要原因是使用了多处理器,以及把数据库直接映射进内存等方法。同时处理数据使用多线程模型。

Mascot 没有使用数据库中解释信息,虽然对有良好解释信息的数据库(如 Swiss-Prot)来说,解释信息对于了解翻译后修饰(post-translational modification)有用,但是,一些从核酸库中直接翻译过来的氨基酸序列往往不包含解释信息。

对于数据库中含有较高重复片段的蛋白质,打分方法无法较好处理。例如,蛋白质 AF005273,虽然分子质量为 1.2×10^6Da,但是超过 80% 的序列是由相同的一段 7 000Da重复片段重复组成,如果所给的实验数据中有一组与这 7 000Da 相符合,则在查询结果中,AF005273 的分值会非常高,而实际上很可能实验蛋白质只包含一段 7 000Da序列。

(二)用串联质谱数据搜索数据库

用串联质谱数据搜索数据库识别蛋白质适用于以下情况:识别复杂的蛋白质混合物;验证肽质量指纹搜索的结果。其流程如图 22.2 所示。

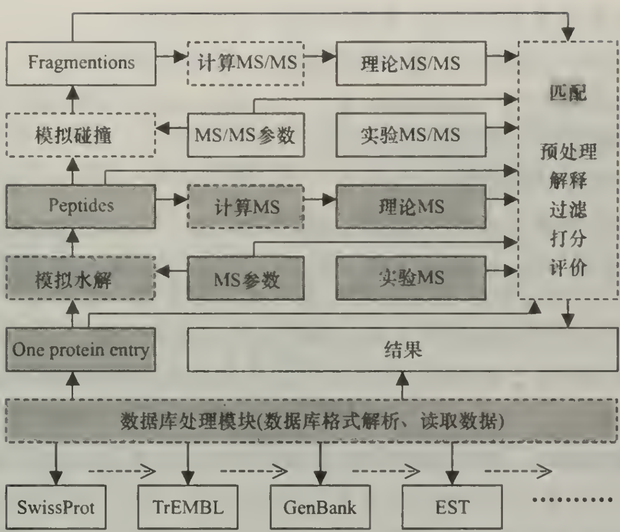


图 22.2　用质谱数据搜索数据库鉴定蛋白质

图 22.2 中实线框是数据,虚线框是计算模块,阴影框是与用 MS 搜索数据库的方法相重合的部分。匹配是核心计算模块,其余大部分是对生物过程的模拟。

数据库处理模块依次从指定数据库中读出每条蛋白质记录。根据酶、遗漏切点等参数模拟水解这个蛋白质,生成所有可能的理论肽。从这些理论肽可以计算出理论 MS。根据实验参数和肽的物理化学性质对每个理论肽可以进行模拟碎裂过程,生成理论碎片离子。从这些理论碎片离子可以计算出每个肽序列对应的理论 MS/MS。

在理论 MS/MS 和实验 MS/MS 比较打分之前,需要对实验质谱信息进行预处理和解释。预处理是为了去除实验质谱中的噪声,提取有效数据。对质谱的解释可以得到序列标签等信息。用这些解释信息和用户输入的一些参数可以直接过滤掉不满足要求的蛋白质序列或肽序列。最后,将理论 MS/MS 与实验 MS/MS 进行比较,给肽和蛋白质打分。这时也可能会用到 MS、序列标签等信息。得到的结果按分值排列,高通量的系统还应对分值的有效性进行评价。显示给用户的结果中可能还包括更多的信息,需要从数据库中读取。

这是所有用 MS/MS 搜索数据库识别蛋白质的系统总体的流程,每个系统侧重点和算法各不相同。但所有搜索数据库的系统都强调打分模块,这也是本节介绍的重点。

1. SEQUEST 算法和系统

SEQUEST 算法和系统[2]是由美国华盛顿大学的 John Yate Ⅲ 开发的,并由 Finnigan 质谱公司整合到质谱数据分析系统并得到进一步优化,已成为目前最常用的蛋白质鉴定软件之一。用 MS/MS 搜索数据库的过程分为四步。

(1) 对串联质谱数据进行处理

① 把碎片离子质核比转换成整数,这样可以将计算速度提高两倍。

②消除母离子两侧 10u 范围的数据,以避免计算出的碎片离子质核比与母离子发生匹配,导致错误结果。

③只保留 200 个最丰富的离子，并把丰值归一化到 100。

④相差在正负 1u 之间的碎片离子丰度被统一到其中较高的值。

⑤标记出质核比较小的 immonium 离子（ + NH_2 = CH(R)），这些粒子经常指示着 His、Met、Trp、Tyr 或者 Phe 等氨基酸的出现。

(2) 搜索与母离子相匹配的序列

搜索数据库中的所有子串，不考虑蛋白酶的特异性。Mass tolerance 在正负 0.05%或者最小 1u 之间。发现 Mass tolerance 在正负 3u 之内都不会影响准确度。质核比计算方法如下：

$$b = \sum a_n + 1$$

$$y = \mathrm{MW} - \sum a_n$$

这时要考虑氨基酸的修饰。有些修饰影响肽的碎裂方式，如 glycosylation；有些修饰是可变的。

(3) MS/MS 匹配的打分方法

与母离子质核比相匹配的序列被模拟打碎，然后与实验 MS/MS 比较，打分时要考虑以下三点：

① 在正负 1u 误差内，与质谱中观察到的离子相匹配的预测碎片离子的数目 n_{i}，以及它们的丰度 i_{m}。

② 若发生连续的离子系列匹配，则加入分数增长项 â(0.075)。

③ 对于质谱中的 immonium 离子(与 His、Met、Trp、Tyr 或者 Phe 等氨基酸相联系)，视序列中是否出现这些氨基酸，加减分数项 $\widetilde{n}$(0.15)。最后，

$$S_{\mathrm{P}} = (\sum i_{\mathrm{m}}) n_{\mathrm{i}} (1 + \beta)(1 + \rho) / n_{\mathrm{t}}$$

(4) 交叉-相关(cross-correlation)分析

上面的打分方法对小质量肽不利，因此 SEQUEST 对用上面打分方法匹配出的前 500 个结果再进行交叉-相关分析。

首先按一定规则对匹配的氨基酸序列重建其质谱，再对实验质谱做适当处理，以使两个质谱之间交叉-相关分析能够反映出碎片离子的相似度。

两个连续信号 $x(t)$和 $y(t)$之间的交叉-相关计算公式如下：

$$C_{xy} = \int_{-x}^{+x} x(t) y(t + \tau) dt$$

其中，τ 是两个信号间的位移值。相关函数实际上是测量了两个信号间的相合度。如果两个信号是相同的，则相关函数在 $\tau = 0$ 处取最大值。离散形式为：

$$R_\tau = \sum_{i=0}^{n-1} x[i] y[i + \tau]$$

其中，x_{i}是重建质谱，y_{i}是实验质谱。两个实函数 x 和 y 的离散相关是离散傅里叶变换对

$$R_\tau \leftrightarrow X_\tau Y_\tau^*$$

的一项。其中，X 和 Y 分别是 x 和 y 的离散傅里叶变换，Y^* 表示复共轭。

交叉-相关的计算方式为：首先从 0～4 096 对两个数据集做快速傅里叶变换，再把一个变换与另一个变换的复共轭相乘，然后对乘积做反傅里叶变换。分数 C_{n}如下计算：

$$C_n = R_\tau(\tau = 0) - \frac{1}{149}\sum_{-75<k<75} R_\tau(\tau = k)$$

最后,分数被归一化到 1.0。

实验发现,排在第一和第二位的氨基酸序列的交叉-相关参量在 0.1 之上时,排在第一位的通常是正确答案。交叉-相关参量改善了对很多小质量肽的打分。

2. Sonar 算法和系统

Sonar MS/MS[3]是利用串联质谱仪鉴定蛋白质的新的软件工具,它包括质量控制评分(quality control scoring)、期望值和一种崭新的 MS/MS 数据图形传递手段。m/z 是第一个真正的蛋白质和多肽质谱数据分析自动化工具,其新的基于信噪的峰发现程序允许其应用于强或弱信号,而用户不必干预。

关于打分方法,把实验质谱的 m/z 轴分成 n 段,每个段取 intensity 的积分,这样就把质谱转换成了 n 维肽空间向量。用同样的方法表示理论质谱,每个离子的 intensity 可是 1,也可以根据质谱仪确定每种离子的强度。如果两个肽向量是平行的,那么它们是相同的。向量的点积恰好反映了它们平行的程度,因此可作为肽匹配的分值。把蛋白质的所有肽向量与质谱向量的点积求和,作为蛋白质的分数。这种方法实质上与 SEQUEST 的交叉-相关类似。

对分数的评价是独立于打分系统的。收集每次搜索的所有分数,并分析这些分值的统计规律。

以上是一次 Sonar MS/MS 识别的高值分数的分布情况。y 轴对应具有特定分数 p(x 轴)的序列的数量。显然,有两类分数:属于左边分布的分数和不属于左边分布的分数。左边的分布对应于随机分数的分布,这是因为当搜索任何大数据库时,大部分的分数是由随机匹配引起的。不属于这个分布的分数是可能的正确识别。

针对分数分布的高值部分,可以把这部分的函数形式拟合到一个 log 函数,即分数与结果数目的对数值。实际中,做这一拟合时,采用了最小方差的方法确定线性系数。这部分分布的可预测性使得可以估计有多少个结果会随机地具有某一分数值 p。对于对数值为 0(大约 $p=47$),就有理由推测,对于任何一次试验,至少会有 1 个蛋白质(或基因)的分数大于 47。期望会具有分数 p 的蛋白质(或基因)的数目在统计学中通常被称为“期望值”(expectation value)。

在上例中,最高的分数是 47 和 50,下一个是 40。分数 50 有效吗?从曲线中计算出的期望值大约是 0.75,这意味着可期望得到 3/4 个结果,或者说在 4 次实验中得到 3 次具有这一分值的结果。显然,这不是个有效结果。

3. Scope 算法和系统

Scope[4]是 Celera 公司设计的打分算法。Scope 算法认为,以往的打分函数没有以结构化和量化的方式利用先验规则,没有明确的模拟仪器测量误差,没有明确的模拟噪声峰值。Scope 算法利用两步随机过程模拟 MS/MS 谱生成过程,解决了上述问题:根据从很多训练样本估计得到的概率分布,生成肽的碎片;根据仪器测量误差,从碎片生成质谱。

(1) 谱生成的随机模型

MS/MS 谱 S:正实数向量;

肽 p:n 个氨基酸残基的序列;

碎片空间:肽 p 生成的所有可能的碎片质核比的集合;

碎裂模式:碎片空间的一个子集;

碎裂空间:所有碎裂模式的集合;

噪声:碎裂模式中不存在的 S 中的峰值;

碎裂:实验条件和肽的物理化学性质决定了观察到一个特定碎片的概率;

测量:由于仪器误差,根据正态分布模拟质峰质核比从它的真正值的偏离。

(2) 打分

$$p^* = \arg\max \psi(S|p),$$

其中 $\psi(S|p)$是给定肽 p,生成质谱 S 的概率密度函数。

$$\psi(S \mid p) = \sum_F \psi(S \mid F, p)\Pr(F \mid p),$$

其中 F 是碎裂模式。

(3) 碎裂概率估计(计算 $\Pr(F|p)$)

学习框架:利用聚类、数据挖掘技术从标注的 MS/MS 数据中学习到肽的特殊性质,这些性质对特定的碎裂事件发生的可能性产生重大影响。

但是,没有这样的学习样本。所以,不得不请教有经验的质谱操作员来选取概率。

(4) 碎片生成质谱(计算 $\psi(S|F,p)$)

做如下假设:

① $m/z = m/z(f)$是一一对应函数;

② $S = S(m/z(f))$是一一对应函数。

其中,第一个假设可以取消。

另外,为每个质峰增加一个噪声碎片,其质核比与质峰相同。

在这些假设下,有

$$\psi(S \mid F, p) = \psi(S \mid \cap_{i=1}^{k}[S_i \overset{M}{=} f_i], F, p)$$

$$\psi(S \mid \cap_{i=1}^{k}[S_i \overset{M}{=} f_i], F, p)$$

$$= \psi(S_1 \mid \cap_{i=1}^{k}[S_i \overset{M}{=} f_i] F, p) \times$$

$$\Pi_{j=2}^{k}\psi(S_j \mid S_1, \cdots, S_{j-1}, \cap_{i=1}^{k}[S_i \overset{M}{=} f_i], F, p)$$

$$= \psi(S_1 \mid [S_1 \overset{M}{=} f_1], F, p) \times$$

$$\Pi_{j=2}^{k}\psi(S_j \mid S_{j-1}, [S_j \overset{M}{=} f_j], F, p).$$

$$\psi(S_1 \mid [S_1 \overset{M}{=} f_1], F, p) = \rho f_1(S_1)$$

$$\psi(S_j \mid S_{j-1}, [S_j \overset{M}{=} f_j], F, p)$$

$$= \begin{cases} \dfrac{\rho f_j(S_j)}{f_{S > S_{j-1}} \rho f_j(S)}, S_j > S_{j-1}; \\ 0, \end{cases}$$

(5) 计算 $\psi(S|p)$

再假设，对于碎片 f，观察到它的概率独立于观察到其他碎片的事件。这个假设不总是成立的，但是没有这个假设就不能有效计算 $\psi(S|p)$。

$$\Phi(i,j) = \sum_{F \subseteq F_j(p)} \psi(S_1, S_2, \cdots, S_i \mid F, p) \Pr(F \mid p)$$

$$\Phi(i,j) = \begin{cases} 1, & \text{if} i = 0, \\ 0, & \text{if} i > j, \\ \Phi(i-1, j-1) & \\ \times \psi(S_i \mid S_{i-1}, S_i \overset{M}{=} f_j) & \\ \times \Pr(f_j \mid p) & \\ + \Phi(i, j-1) \Pr(\overline{f_j} \mid p), \text{otherwise}. & \end{cases}$$

$$\psi'(S \mid p) = \max_{F \subseteq F(p)} \psi(S \mid F, p) \Pr(F \mid p)$$

这两个值非常相近，计算后者的好处是可以取中值的对数，以避免计算 $\phi(k,m)$ 时的数字不稳定性。

(6) 实现细节

① 很多碎片质量相同。事件"这个质量出现"是事件"所有具有这个质量的碎片都没出现"的补。

② 算法的好坏取决于假定的肽碎片概率、被测量到的质核比对于真实值的分布形式。在没有训练数据的情况下，只能靠人工估计。

③ 函数 $\psi(S|p)$ 是连续概率密度函数，所以不适合用作打分函数，因为任何特定质谱的实现都是零概率事件，必须决定如何用 $\psi(S|p)$ 给肽打分。并不是把质谱当作实数向量，而是认为每个谱峰有一个小的宽度，这个宽度与质谱仪测量质核比的精度有关。这样就能计算质谱是由一个特定肽的碎片产生的概率 p。分数等于 $-\log(p)$。

4. Mascot 系统

Mascot 对 MS/MS 数据的预处理包括以下几步：

(1) 取消了对数据库的索引处理过程。

(2) 提供了对蛋白质质量的限定搜索。不过把蛋白质质量作为一个移动窗口，对数据库中质量大于输入质量的蛋白质序列进行处理。比如，指定蛋白质质量为 20×10^3Da，那么，数据库中每一个超出 20×10^3Da 的蛋白质，程序在 20×10^3Da 窗口范围内对序列打分，得出最高的匹配。这样可以避免蛋白质质量比数据库中该蛋白质条目的质量小的情况。例如，Swiss-Prot 的条目 INS-BOVIN 蛋白质包含了信号和连接的肽，所有累计为 11 394Da，而实验室测得的分子质量为 5 374Da。

(3) Mascot 也利用峰的丰度值。实验表明，可以认为低于某个丰度以下的所有峰都是随机噪声，这样处理后的数据提交给程序用于对肽质量搜索影响不太大。其处理步骤为：

① 每隔 14Da 取一个丰度最高的峰，这些峰全部保留。

② 剩余的峰处理：确定某个丰度值，对该丰度值以上的峰保留，以下的丢弃。

③ 上述生成的所有的峰进行规格化。

(4) 提供物种信息过滤。

Mascot 的打分方法是：假设某一匹配是随机匹配的概率为 p，那么分数为 $-\log(p)$。

5．其他系统

MS-Tag[5]是 ProteinProspector 软件包中的一个。MS-Tag 有两种工作方式，一种是利用某个肽的质谱数据进行数据库查询，得出数据库中可能与这个质谱数据匹配的所有肽的序列。另一种是结合 *de novo* 序列解释方法：计算出所有可能的氨基酸组成；为每种组合在线地生成所有序列，再由 MS-Tag 逐个检查，就像查找一个所有可能组合的数据库。这种方法缺点很明显：数据需要很好的质量精确度和 immonium 离子表示；随着肽序列增长，排列数呈指数增加，所以此方法要求肽小于 1 300Da。

MS-Tag 不使用打分系统。其理论是：MS/MS 质谱数据经过预处理后，将每一个 MS/MS 质谱数据都看作是一个碎裂离子数据，对蛋白质数据库中的每一个肽，测试它所能生成的所有离子类型，看是否与某个 MS/MS 质谱数据匹配。一个肽序列所能生成的碎裂离子与输入的 MS/MS 质谱数据匹配得越多，它越可能是查询结果。然后按下列规则将多个匹配排序：①生成较少的不匹配的碎裂离子的肽序列排前；②按肽质量由小到大排序；③按索引值由小到大排序。

PepSea[6]使用肽序列标签搜索数据库。肽序列标记是蛋白质水解后，肽的一小段氨基酸序列加上该肽的质量以及肽碎片的质量信息。这种方法速度很快，不过需要在搜索之前提取肽序列标记(虽然这也可以自动化地提取)。

PepFrag[7]使用与其他质谱信息(如氨基酸组成成分)相结合的肽片段质量信息鉴定蛋白质。

（三）对各种算法和系统的评价

不管是 *de novo* 测序，还是数据库搜索方法，都需要对输入的实验质谱数据进行预处理，以减少噪声，提取有效数据。对于质峰的选取，需要确定选取多少数目的质峰和哪些质峰。简单的方法是选取丰度值最高的固定数目质峰，如 SEQUEST。而 Mascot 则动态地确定要选取的质峰的数目，它考虑了匹配时肽的性质(如长度)。更复杂的去噪声的办法还有利用最大熵原理，如 Micromass[8]等。但仅凭丰度值选取质峰是不够的，因为有用的质峰可能被噪声埋没。所以，合理的做法是加入先验知识。Mascot 通过在每一个固定范围内(14Da)保留一个质峰的方法加强了对小丰度值质峰的选取，以获得与离子系列相对应的质峰阶梯。也有人考虑了峰之间的相关性来调整质峰的丰度，以选择最优的一组峰，如 AuDeNS[9]。笔者认为，充分利用关于质谱的先验知识选取质峰应是未来的趋势，理论完备的信号处理方法未必适合解决这一特殊的生物问题。

在识别蛋白质时，完全依赖于对质谱的解释就成了 *de novo* 的方法。搜索数据库的各个系统也或多或少地加入了解释模块，其目的是直接过滤掉不符合的序列以加快搜索速度和减少错误匹配。最简单的方式就是利用蛋白质或肽的质量信息进行过滤，如 SEQUEST 和 Mascot 等。PepFrag 利用了氨基酸组成信息(氨基酸组成与氨基酸序列是不同的概念，氨基酸组成只是说明氨基酸的出现，没有位置信息)。有的还加入蛋白质的等电点值，如 MS-Tag 等。更复杂的是从质谱中解释出部分氨基酸序列信息，称作序列标

签，如 PepSea。准确的序列标签对于消除假阳性非常有效。值得注意的是，过多的和不正确的解释也可能误导搜索的过程。多大程度的解释才算合适，以得到最好的搜索效果是有待探讨的问题。无论如何，有效的解释模块对于提高系统的性能是非常重要的，这也是未来搜索数据库系统的发展趋势。

打分算法是搜索数据库算法的核心。普通的打分算法包括简单地对匹配的质峰计数，累加匹配的质峰的丰度值，或者把丰度值取对数后再累加。取对数的目的是弱化丰度的贡献。SEQUEST 除了累加匹配的峰的丰度值外，还考虑了特定信息对分值的加权。SEQUEST 还采用交叉-相关分析重新对用上述方法得出的前 500 个结果重新打分，以加强对小肽的打分。Mascot 是基于概率的打分算法，提出了 *P* 值的概念，即计算匹配为随机的概率，再由这个概率计算分数。Sonar MS/MS 把质谱转化为向量，用向量的点积作为分值，这实质上也是交叉-相关分析。SCOPE 利用 Bayes 模型进行打分，即对于给定的质谱求每个序列的后验概率作为分值。这些打分算法复杂程度各异，也各有优缺点。交叉-相关分析来源于信号处理理论，用于比较两个信号是否一致，但未必适合于质谱这一特殊信号。如果简单地把质谱当作普通信号进行处理，不加入先验知识，那么必将丢失很多信息。Mascot 的 *P* 值应该说是对匹配很好的一个测量，文献中没有给出计算方法。但这样的 *P* 值必定是不可计算的，所以只能是个估计值，仍要做很多近似和利用经验信息。SCOPE 打分算法的理论模型完备，但为了便于计算已做了很多假设和近似。同时它还没有利用丰度信息，对噪声的模拟也没什么道理。笔者认为，融入更多的专家经验，充分利用对质谱的解释信息，考虑更多的因素比好看的理论模型更重要。打分方法仍是个尚未完全解决的问题。

高通量的蛋白质识别应尽量排除人的干预。所以，在打分之后应对分值的有效性进行评价，也就是要确定匹配是完全随机的还是真实匹配。SEQUEST 在进行交叉-相关分析之后，凭经验发现如果排在第一位和第二位的结果的分值相差在 0.1 以上，那么排在第一位的结果通常是真实匹配。这只是个特殊情况。一般的做法有三种：一是使用模拟的方法，首先在肽指纹搜索中提出[10]。在模拟中，随机地构造理论质谱，再用理论质谱搜索数据，最高的分数被保存。反复进行这一过程，就可以得到一个随机识别的分数分布。从这一分布就可以计算实际识别的蛋白质或肽为随机的概率。另一种方法是直接计算随机匹配的概率。比如 Mascot 认为 *P* 值在 0.05 以下的匹配才是真实的。然而，这些直接计算的方法不如模拟的方法可靠，因为计算的过程非常复杂，不得不做近似处理。第三种方法是在每次搜索中统计分数的分布。例如，Sonar MS/MS 提出了期望值的概念，期望值在 1 以下的匹配才被认为是有效的。一个统一的评价标准对于比较各个系统的性能是必需的。

总之，如何更有效地提取质峰、对质谱做更合理的解释、设计灵敏性和选择性更高的打分算法以及对分值的有效性做更准确的评价，仍是用质谱搜索数据库的方法进行高通量蛋白质识别研究中需面临的挑战。

（四）从串联质谱数据直接推测肽序列（*de novo* 测序方法）

1. 原理和算法

蛋白质鉴定的一个重要方法是使用一个蛋白质数据库来鉴定肽：给定肽的串联质谱，首先从数据库中选出与其母肽质量相匹配的那些肽（作为候选肽），然后把给定的质谱与

数据库中的候选肽的理论质谱相比较，最后，将这些候选肽按照它们的质谱与给定肽的质谱的匹配程度进行打分排序。但这种方法也有缺点：它们要依赖一个蛋白质数据库以及该数据库中蛋白质条目的正确性；普遍存在的翻译后修饰会降低肽鉴定的精度；随着数据库中蛋白质序列数的增长，数据库搜索的鉴定方法会越来越慢；这种方法不会发现新蛋白质。即使 Sequest 等软件系统考虑了某些翻译后修饰(post-translational modifications)，未测序的蛋白质以及作为选择性剪接(alternative splicing)产物的蛋白质也还是没有被考虑进来。在此情况下，就特别需要不使用来自蛋白质数据库的信息而直接解释串联质谱数据的技术(但此时要使用 20 种氨基酸残基的质量等信息)。这种直接解释串联质谱数据的蛋白质鉴定方法就称为蛋白质(或肽)*de novo* (从头)测序。

通常情况下，肽 *de novo* 测序分两步：第一步，生成与给定串联质谱数据相匹配的理论肽的集合。这里需要说明的是，由于不可避免的污染和测量误差存在，本集合可能非常大(元素可能上万个)。第二步，使用启发式方法对这些理论肽进行评分排序，然后输出评分最高的那些肽。

与自动化的相关数据库搜索技术不同，使用质谱数据从头确定肽的氨基酸序列是可能的。这种方法的特点就是要有专业质谱分析人员对有关数据做出人工解释。使用质谱检测、以自动化的方式进行肽的从头测序分析，这个问题目前还未彻底解决。然而，为解决此问题，正处于研究之中的一些算法，能够用来确认以自动化、高通量的方式搜索相关数据库所获得的结果。SHERENGA[11]就是一个这样的算法，它能从串联质谱数据产生一个可能的肽序列排序表。一个类似的方法[12]基于这样一个前提——如果所确定的 *de novo* 序列(正确或错误)能够比数据库导出的序列更好地说明 MS/MS 数据，那么从数据库搜索中得到的序列可能是假阳性；另一方面，如果数据库搜索导出的序列比 *de novo* 序列更好或者可同样说明 MS/MS 数据，那么该导出序列可能是一个适当的匹配。在其他途径中，为了从同一个数据集中自动导出 *de novo* 序列，使用肽的甲基醇化作用在两个蛋白质混合物之间相对量化蛋白质，文献[13]研究了质谱方法的能力。

对于肽 *de novo* 测序，一般要求质谱数据的质量较高，肽断裂情况良好，而且氨基酸不能有修饰，其技术难点在于串联质谱数据的预处理和如何重新标定质谱。这种方法的优点是即使不能完全测对，也能提供重要的肽序列标记(几个氨基酸组成的短肽)，供数据库搜索参考。然而，这种方法也有不足之处，就是不管方法本身的容错能力多强，如果重要的离子质谱未测出来，那么所提供肽序列的可信度也要大打折扣。从目前实际应用看，考虑到数据库搜索方法和肽 *de novo* 测序方法各有优缺点，肽 *de novo* 测序与数据库搜索相结合的混合蛋白质鉴定方法是首选。

已有一些人，如 Fernandez-de-Cossio 等[14]，Taylor 和 Johnson[12]，以及 Dancik[11]等人提出了几种效率不等的生成匹配肽的集合算法，他们都是将质谱转化成有向无环图(directed acyclic graph)，其中，图中每一个连通的路径表示可能的肽序列，也就是将肽测序问题转化为在有向无环图中寻找反对称最长可行(最佳匹配)路径问题(antisymmetric longest path problem in directed acyclic graph)，如图 22.2。他们使用了不同的算法在多种可能路径中选择好的匹配序列。已有几个软件包，如，Lutefisk[12,15,16]，SHERENGA[11]，AuDeNS[17]以及其他系统[18~22]实现了这种 *de novo* 肽测算法。

质谱图中的顶点对应于质谱图像的质峰；如果某一对顶点对应的质谱峰值之差为某

个氨基酸(粗略地说)的质量,则表示这一对顶点之间可以建立一个有向边,并且用这个氨基酸标记该边。

2000年,Chen等人[23]提出了一个使用动态规划的 *de novo* 测序算法。该算法有两种形式:一种针对干净(无噪声等干扰)的数据;另一种针对含有噪声的数据。因为干净的数据在生物实验中是不存在的,所以,针对噪声数据的 *de novo* 测序算法才是可行的。Chen等人证明了针对噪声数据的 *de novo* 测序算法的时间复杂度至多为给定质谱数据质峰数的立方。然而,他们并没有在实际系统中给出该算法的具体实现。在计算上直接使用该噪声算法太复杂,因为那样会使可能解的数量太大。另外,也需要考虑测量误差。

最近,也有一些人[24]充分利用有效质谱峰的强度一般都较高的特点,在生成匹配肽的集合算法上,提出了不同于前面提到的图模型,而是直接构造求待测序肽的质谱峰强度之和最大的最优化模型。从文献看,他们实现的软件系统PEAKS[25]的肽 *de novo* 测序优于当前所有的肽 *de novo* 测序软件工具,而且,最近他们已对PEAKS进行了改进,使其成为除具有 *de novo* 测序能力外,还具有数据库搜索能力的混合的串联质谱蛋白质(肽)鉴定软件工具。

2. *de novo* 肽测序软件

(1) 自动化 *de novo* 肽测序软件工具 AuDeNS

Baginsky等人[17]提出了对输入的质谱峰赋予相关性值的启发式的数据清洗算法(割草机算法,"grass mowers"),而且实现了一个原型系统——AuDeNS。该系统首先使用割草机算法预处理原始质谱数据,然后利用经修正的Chen等[23]的噪声测序算法,该算法能够处理质谱测量误差。而且,他们通过对解赋予相关性值,以及仅列举用户指定的、与最大相关性值有关的域值内的那些解,解决了可能的有关解的数量问题。AuDeNS系统的输出是一列带有评分排序的多个序列(这是由输入数据本身固有的模糊性决定的)。虽然AuDeNS在性能上还未超过 *de novo* 肽鉴定工具Lutefisk,但Baginsky等人相信AuDeNS最终能赶上或超过Lutefisk,因为AuDeNS远比Lutefisk或Sequest运行得快,现有的启发式清洗算法还可进一步微调,而且还可以增加新的清洗算法以及其他可以改进的方法。AuDeNS工作步骤如下:

① 应用一些数据清洗算法处理输入的质谱数据,为每一个质峰 i 赋予一个表示其相关性的值 $r(i)$。

② 应用修正的肽测序算法。Baginsky等人的这个测序算法是基于Chen[23]提出的针对噪声数据的动态规划算法的。Chen[23]的算法是求解所有质峰对(即图模型中的边)的权重之和的最大化问题(即最长路径);而Baginsky等人的修正算法是求解所有质峰被赋予的相关性的值之和的最大化问题。在实验质谱图中,质峰的相关性比质峰的权重更具体,因而后者的算法较前者更具有合理性。

(2) 使用MS/MS数据的肽 *de novo* 测序软件工具 PEAKS

2002年在美国Orlando召开的第50届ASMS会议上,加拿大西安大略大学的Bin Ma、Kaizhong Zhang等人推出了他们开发的功能强大的使用MS/MS数据的肽de novo测序软件工具——PEAKS,该软件能够自动地从肽的MS/MS实验数据导出该肽对应的氨基酸序列。PEAKS接受MS/MS数据的质谱峰列表之后,试图用每一个可能的肽可产生的各种离子解

释质谱中的质峰，最后，把其中解释质峰最好的那些肽作为预测的肽序列输出。

PEAKS 的新颖之处在于这种方法依赖于更多的信息，而且能容忍质谱仪的测量误差。PEAKS 的方法不同于像 MSCOT[26]、Sequest[27]等软件系统所使用的数据库搜索方法。PEAKS 是搜索所有可能的氨基酸串，而不是搜索蛋白质序列数据库。而且为了加快搜索速度，PEAKS 使用了一个复杂的动态规划算法，该算法有别于图模型方法[15,28~30]中常用的动态规划算法。PEAKS 的动态规划算法能保证最好的那些肽不被忽略。一般情况下，PEAKS 使用中等水平的台式计算机能够在 10s 内预测出经胰蛋白酶切得到的肽(由 10～20 个氨基酸组成)的序列。据说[24]，就精度而言，PEAKS 的预测能力超过了所有的现阶段能见到的其他肽 *de novo* 测序软件工具。PEAKS 处理质谱的步骤如下。

① 质谱数据的预处理。包括噪声过滤、去卷积以及将双电荷、三电荷离子转换为单电荷离子。对于噪声过滤，先要估计谱的噪声水平 h_0，然后将所有质峰的强度减少 h_0，最后删除峰强度小于或等于 0 的质峰；对于去卷积，则先要检测出包含高强度质峰的邻接序列的窗口，然后对每一个窗口，将其中的所有质峰求和作为本窗口“中心”的质峰强度，最后移去除了该“中心”以外的所有质峰；对于双电荷、三电荷离子，先要使用同位素离子确定离子所带电荷数，然后把多电荷质峰转换为单电荷质峰，再把所有的同位素离子峰强度加到它们所对应的单一同位素离子峰上，最后移去同位素离子峰。

② 候选肽的计算。首先定义质峰与离子的匹配评分，然后再定义肽与质谱数据(即由一列质峰数据构成)的匹配评分。一个质峰与离子的匹配评分定义为该离子对应的质峰强度的对数值乘以与该离子和质峰质量之差有关的一个因子；而肽与质谱数据的匹配评分则定义为所有质峰与各种离子匹配评分之和，这些离子包括：a、b、c、x、y、z 型离子，以及 y、b 型离子失水、氨后得到的离子。如果某个峰与两个不同的离子匹配，只有匹配评分较大的才被计算在内。有了肽与质谱数据的匹配评分的定义后，在此，肽 *de novo* 测序问题则可定义为：给定一个观测质谱数据，查找与之对应的匹配评分最大的肽。最后应用一个复杂的动态规划算法计算出与给定质谱数据匹配得最好的肽。通常可选取评分最好的前 10 000 个肽序列作为候选肽序列。

③ 候选肽的排序。通过更加精确的评分计算，对第二步中得到的候选肽进一步评估。在本步中，要考虑到 immonium 离子和内部分裂离子。离子质量允许误差也要比第二步的减小。此外，质谱数据也要重新标度，以体现由于温度变化引起的质谱仪误差。通过使 y 离子的 m/z 值与对应的质峰相适应，计算出重新标度的因子。本重新标度方法极似 Taylor等[12]使用的方法。待所有的候选肽的新评分计算完之后，具有最高评分的前 n 个肽序列被输出，其中 n 为用户指定的正整数值。

3. 将来可能的发展方向与趋势

一方面，肽 *de novo* 测序问题本身可从以下方面继续研究：①充分考虑实际质谱实验数据特点，研究更好的串联质谱数据预处理方法；②构造与预处理方法相适应的更加合理的模型、算法；③利用统计方法对已知的蛋白质序列数据进行分析，进一步处理上一步输出数据。

另一方面，由于现阶段肽 *de novo* 测序问题针对的原始数据的局限性(即串联质谱实验数据与理论数据相比还不够理想)，研究肽 *de novo* 测序方法与数据库搜索方法相结合

的蛋白质鉴定的混合方法，或许更实用一些。

（五）高通量蛋白质识别

对 PMF 进行操作的过程完全可以依靠计算机程序化，即对高通量的蛋白质实现自动识别。这也是实际的需要，通常在实验中，从质谱测得的各多肽的分子质量的数据是批量的，当对数百个蛋白质进行识别时，则可能产生数千个多肽分子质量，这样的工作量是不可能通过人工来完成的。必须建立一个自动进行的高通量蛋白质识别系统（protein identification pipeline）。

在为高通量的蛋白质识别建立一个完整的计算流程（pipeline）之前，有必要明确所建立的这个系统所能够处理的任务的范围，计算资源的消耗以及复杂度。对于一个经济的计算资源的消耗，在进程的每一个阶段必须具有同等的通量，必须识别且排除瓶颈的出现，顺利而有效地处理从 MS 获得的大量数据。

首先必须确定对于高通量蛋白质识别的 MS 数据处理的具体参数。例如，工作量、日期、时间、查询类型、状态等，以及返回数据库的查询结果应如何分析。作为一个高效的识别系统，应该可以根据结果来确定后面分析的回路。

三、对实验结果的分析

一旦蛋白质的序列信息明了，下一步就是通过各种分析工具了解该蛋白质的性质，如等电点、分子质量、同源性、亚细胞定位等，以此成为蛋白质功能分析的基础。目前这方面的工具主要有十多种（http://www.expasy.org/tools/）。本节将从蛋白质的序列渐进到其结构和功能，逐一介绍应用于分析和预测蛋白质的生物信息学软件。

（一）蛋白质序列的生物信息学分析

蛋白质序列的生物信息学分析是从理论到实践、最初也是最重要的部分，对于实验思路的制定可以提供极其重要的指导信息。下面将逐步介绍分析蛋白质序列的过程和软件。

1. 序列检索

从质谱等方法鉴定蛋白质获得的序列可能不完整，或是蛋白质家族中的一种，这时需要用蛋白质名称进行序列检索，获得蛋白质的全序列。表 22.1 列出了主要的蛋白质序列检索工具。

表 22.1 主要的蛋白质序列检索工具

软件	网址	概述
Entrez	http://www.ncbi.nlm.nih.gov/Entrez	由 NCBI 提供的集成检索工具，可以在网络上运行，也可以本地化
DBGET	http://www.genome.ad.jp/dbget	日本京都大学化学研究所提供的检索 GenomeNet 的数据库检索工具
SRS	http://bioinfo.biosino.org/srs6/或者 http://srs.ebi.ac.uk/	由 EMBnet 提供的主要数据库检索工具

2. 与核酸相关的软件

如果想了解编码这段蛋白质的核酸序列，ExPasy 上提供了很多 DNA-蛋白质以及蛋白质-DNA 之间相互转换的工具，见表 22.2。

表 22.2 Protein-DNA 之间相互转换的工具

软件	网址	概述
Translate	http://us.expasy.org/tools/dna.html	核酸序列转换为蛋白质序列
Transeq	http://www.ebi.ac.uk/emboss/transeq/	Emboss 软件包中的转换工具
Graphical Codon Usage Analyser	http://gcua.schoedl.de/	以图形的方式表示密码子偏好性
BCM search launcher	http://searchlauncher.bcm.tmc.edu/seq-util/Options/sixframe.html	以 6 个读码框来翻译一段序列
Backtranslation	http://www.entelechon.com/eng/backtranslation.html	一段蛋白质序列反翻译为核酸序列
FSED	http://ir2lcb.cnrs-mrs.fr/d-fsed/fsed.html	探测移码错误
LabOnWeb	http://www.labonweb.com/	分析 est 的工具
GeneWise	http://www.sanger.ac.uk/Software/Wise2/genewiseform.shtml	蛋白质序列和基因组序列进行比对，允许内含子和移码错误

3. 蛋白质基本物化性质分析

蛋白质全序列获得后，接下来就是分析蛋白质的基本特性，即物化特性。主要包括分子质量、理论等电点、氨基酸组成、原子组成、消光系数、估计半衰期、稳定指数、亲(疏)水性等。已开发了许多有用的计算工具用于分析蛋白质的基本性质。其中大部分可以从 Geneva 大学医院和 Geneva 大学的 Expasy 站点获得，见表 22.3。

表 22.3 蛋白质基本物化性质分析

软件	网址	概述
AACompIdent	http://us.expasy.org/tools/aacomp/	以氨基酸的组成在 Swiss-Prot 里面寻找与其类似的蛋白质
AAComSim	http://us.expasy.org/tools/aacsim/	把 Swiss-Prot 中某一条蛋白质与里面其他的比较，给出与其氨基酸组成非常接近的蛋白质
MultiIdent	http://us.expasy.org/tools/multiident/	用等电点、分子质量、氨基酸组成、序列标签和指纹图谱这些指标来鉴定蛋白质
PeptIdent	http://us.expasy.org/tools/peptident.html	通过等电点、分子质量和指纹图谱来鉴定蛋白质
TagIdent	http://us.expasy.org/tools/tagident.html	给出一系列的与给定等电点，分子质量相近的蛋白质，也可以通过很短的序列标签和指纹图谱来帮你鉴定蛋白质
FindPept	http://us.expasy.org/tools/findpept.html	通过实验数据鉴定非特异性剪切下来的肽段
PeptideMass	http://us.expasy.org/tools/peptide-mass.html	计算 Swiss-Prot 或者 TrEMBL 中某一条蛋白质的分子质量或者用户自定义的蛋白质
CombSearch	http://us.expasy.org/tools/CombSearch/	一个整合了多种网络鉴定方法的工具
PepSea	http://195.41.108.38/PepSeaIntro.html	通过序列来鉴定蛋白质
ProtParam	http://us.expasy.org/tools/protparam.html	计算多种理化指标
Compute pI/MW	http://us.expasy.org/tools/pi-tool.html	计算等电点和分子质量
ProtScal	http://us.expasy.org/cgi-bin/protscale.pl	氨基酸度量表示
MW, pI, Titration curve	http://www.up.univ-mrs.fr/~wabim/d-abim/compo-p.html	计算等电点、分子质量，并以曲线的形式表现出来

（二）序列特性分析

蛋白质序列特性对于蛋白质高级结构以及其功能来说具有决定性的作用，所以，一级结构的分析是功能预测和结构预测的基础，了解蛋白质序列中的特殊性是非常重要而且是必要的。通过相应的软件分析序列可以得到大量详尽的信息。

1. 跨膜区序列分析

蛋白质序列含有跨膜区域提示它可能作为膜受体来发挥作用，也可能定位于膜的离子通道蛋白质。这种蛋白质往往和细胞的功能状态密切相关。目前有许多种预测跨膜螺旋的方法。当然通常都是观察疏水氨基酸的分布区域来预测的。还有其他的技术主要是对已知跨膜螺旋的研究得到的。目前跨膜螺旋的数据库最重要的是 Tmbase 数据库。这些技术预测的准确性可以达到 80%～90%。以 TMPRED 为例，它是基于对于 Tmbase 数据库的统计学分析来预测的。

2. 卷曲螺旋分析

卷曲螺旋在很多蛋白质中存在，如亮氨酸拉链结构、肌球蛋白质。α 螺旋可以直接从序列中预测的。如 COILS(http://www.ch.embnet.org/software/COILS_form.html)。

3. 信号肽分析

信号肽通常都是一段疏水氨基酸序列。含有信号肽的蛋白质一般被分泌到胞外，对于研究蛋白质定位有很大的帮助。SignalP 是目前应用比较广泛的软件。具体用法将在蛋白质亚细胞定位一节中具体说明。

表 22.4　蛋白质序列特性分析软件

软件	网址	概述
REP	http://www.embl-heidelberg.de/~andrade/papers/rep/search.html	寻找重复序列
REPRO	http://mathbio.nimr.mrc.ac.uk/~rgeorge/repro/	寻找重复序列
RADAR	http://www.ebi.ac.uk/Radar/	寻找重复序列
SAPS	http://www.isrec.isb-sib.ch/software/SAPS_form.html	对于蛋白质序列的统计应用分析.(EBI)
Coils	http://www.ch.embnet.org/software/COILS_form.html	蛋白质中螺旋区的预测
Paircoil	http://nightingale.lcs.mit.edu/cgi-bin/score	蛋白质中螺旋区的预测
Multicoil	http://nightingale.lcs.mit.edu/cgi-bin/multicoil	蛋白质多重螺旋区的预测
PEST	http://www.icnet.uk/LRITu/projects/pest	PEST 区域的寻找
PESTfind	http://www.at.embnet.org/embnet/tools/bio/PESTfind/	PEST 区域的寻找
HLA-bind	http://bimas.dcrt.nih.gov/molbio/hla_bind/	HLA 肽段序列结合位点的预测
SYFPEITHI	http://syfpeithi.bmiheidelberg.com/scripts/MHCServer.dll/home.htm	和 MHC 相关的预测

4. 其他

对于其他如重复序列、PEST 区域的寻找,等等,可选择 PESTfind[33,34]软件。如果某种胞内蛋白质的半衰期在 2h 以内,通常这个蛋白质序列中就会有一些特殊的区域,这些区域富含脯氨酸,谷氨酸,丝氨酸和苏氨酸(P、E、S 和 T),即 PEST 域。PESTfind 就用来寻找这些特殊的区域。

(三)预测翻译后修饰的软件

翻译后修饰对于蛋白质功能具有决定性的作用,研究这些修饰对于了解蛋白质的功能有很大的帮助。目前软件的工作主要是预测蛋白质翻译后修饰的方式,以及其修饰位点。当然不同的工具能帮助分析不同的修饰方式。例如,FindMod[32]是根据目前已知蛋白质的 22 种翻译后修饰的理论而建立起一种独特的识别模式来进行翻译后修饰的预测。预测蛋白质翻译后修饰软件如表 22.5。

表 22.5　预测翻译后修饰的软件

软件	网址	概述
FindMod	http://us.expasy.org/tools/findmod/	预测蛋白质潜在的翻译后修饰的可能性,并且在肽段中寻找可能被替代的单个氨基酸残基
GlycoMod	http://us.expasy.org/tools/glycomod/	通过实验得到的数据来预测寡糖链发生的可能性
GlycanMass	http://us.expasy.org/tools/glycomod/glycanmass	计算相应寡糖链的分子质量大小
Peptidecutter	http://us.expasy.org/tools/peptidecutter/	预测一段给定蛋白质中潜在的蛋白酶剪切位点
NetOGlyc	http://www.cbs.dtu.dk/services/NetOGlyc/	预测在哺乳动物蛋白质中糖基化 *O*-连接的位点
NetNGlyc	http://www.cbs.dtu.dk/services/NetNGlyc/	预测在哺乳动物蛋白质中糖基化 *N*-连接的位点
DictyOGlyc	http://www.cbs.dtu.dk/services/DictyOGlyc/	预测在网柄菌中 *O*-连接的位点
YinOYang	http://www.cbs.dtu.dk/services/YinOYang/	预测在真核细胞蛋白质序列中 *O*-糖基化的位点
Big-PI Predictor	http://mendel.imp.univie.ac.at/sat/gpi/gpi-server.html	GPI 锚点预测
DGPI	http://129.194.185.165/dgpi/	GPI 锚及其剪切位点的预测
NetPhos	http://www.cbs.dtu.dk/services/NetPhos	在真核细胞蛋白质中预测丝氨酸、苏氨酸、色氨酸的磷酸化位点
NetPicoRNA	http://www.cbs.dtu.dk/services/NetPicoRNA/	预测 RNA 病毒蛋白质中蛋白酶的剪切位点
Sulfinator	http://us.expasy.org/tools/sulfinator/	预测酪氨酸硫化位点
NMT	http://mendel.imp.univie.ac.at/myristate/SUPLpredictor.htm	预测 N 端的烷基化

（四）序列相似性比较

得到蛋白质序列和基本性质的信息后，下一步就是对于蛋白质功能的分析和预测了。一般意义上，如果两条序列具有一定程度上的相似性，那么我们也认为它们会执行相似的功能。

序列相似性比较是分析和预测的基础，也是生物信息学中最频繁的操作。目前最常用的程序是 BLAST 和 FASTA。FASTA 进行的通常都是整体配对比较，特异性较高，但是敏感性较低。现在大多数的研究都是通过 NCBI 开发的 BLAST(basic local alignment search tool)来进行相似性比较和搜索。他们采用 BLAST 的基本思想，分别用特定的序列数据库和特定类型的输入序列来完成不同的工作。在 BLAST 家族中，与蛋白质相关的程序有：BLASTP，在蛋白质序列数据库中比对蛋白质序列；BLASTX，在蛋白质序列数据库中比对该蛋白质的核酸序列(包括 6 种读码框)；TBLASTX，在核酸序列数据库中比对蛋白质序列(包括 6 种读码框)。

本章主要以 NCBI 提供的 BLAST[34~53]为例，说明其具体用法。

在浏览器地址栏中输入 http://www.ncbi.nlm.nih.gov/blast，如图 22.3 所示。

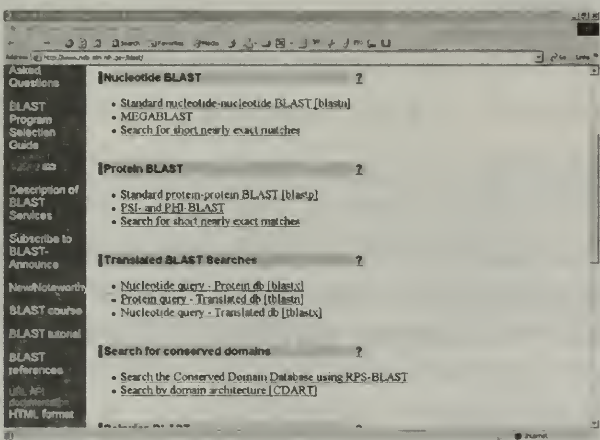


图 22.3　BLAST 界面 1

该网页右侧有 9 项。

第一项：BLASTN，最常规的核酸序列之间的比较。

第二项：BLASTP，标准蛋白质序列比较。其中 PSI-BLAST 通过寻找蛋白质家族中的保守序列进行比对，并对数据库进行多轮循环检索，敏感性相当高。

第三项：包括 BLASTX，TBLASTX，TBLASTN。将序列翻译后再进行比较，以寻找蛋白质同源信息。

第四项：RPS-BLAST，已知功能域的查询。

第五项：PAIRWISE BLAST，序列之间的相互比较。

第六项：基因组比对。将实验得到的基因组序列和已确定测序完成的基因组数据库进行比较。

第七项：其他特殊比对。比如 IGBLAST，分析 Gen Bank 已存入的免疫球蛋白质序

列。

第八项:当比对的时间较长时,NCBI 会提供一个计算 ID 号,在其中输入 ID 号即可得到先前的计算内容。

第九项:针对一些老版本的浏览器不支持的特性的安排。

了解了基本特性后,下面可开始“实战演习”:

选择第二项 BLASTP,如下:

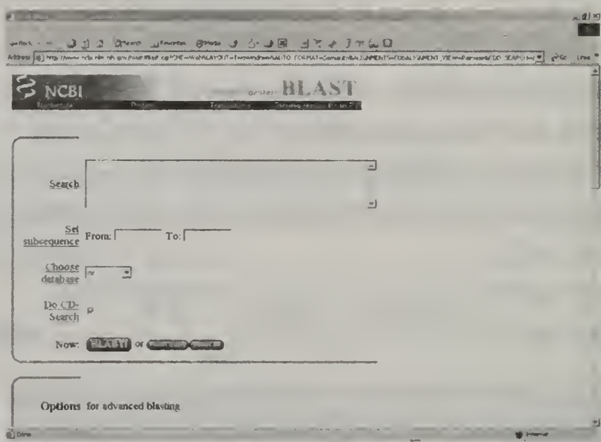


图 22.4　BLAST 界面 2

有三个红色方框区域(图 22.4 中只表示了两个红色方框):

第一个红色方框内示使用缺省参数的 BLAST,在 search 文本框内输入序列后选择你所需比对的数据库,subsequnce 这一项可以不填。如果你只想搜索序列中的某一段,只需在 subsequnce 框内填入序列位置即可。点击 BLAST,之后出现一个网页告诉你刚刚提交的查询在 NCBI 的排队序号,以及大概要等待的时间。耐心等待,点击 FORMAT 可得到刚才的“演习”结果。

BLAST 结果分为四个部分:

第一部分:BLAST 程序本身和所要比对的序列和数据库的信息。

第二部分:图示比对结果。如果将鼠标移至图中表示搜索结果的红色长条上,图形上方文本框内便会出现相应的搜索结果的数字信息。点击,直接出现详细比对信息(即第四部分)。第二部分是第三部分非常直观的图形化显示。

第三部分:比对结果小结。每一行的开头是序列名称,直接链接到 NCBI 里面的序列存储信息。接着是比对得到的分数和期望值,点击,到第四部分相应的详细比对结果。

第四部分:更加详尽的比对结果。

第二个和第三个红色方框是对第一个的补充。具体情况请参考 NCBI 的帮助文档(http://www.ncbi.nlm.nih.gov/Education/BLASTinfo/information3.html)。

BLAST 结果可以给出该序列的详细信息或是其同源序列信息,进而确证或推断出该序列的特征和功能。同时还能链接到 Gen Bank 库中,察看此序列的提交以及物种情况,进一步得到 NCBI medline 的登陆号,查看相应参考文献摘要。提供 BLAST 服务的还有 ExPASy,EBI,PBIL 等,具体网址及其他类似比对程序见表 22.6。

表 22.6　提供 BLAST 服务的网址及其他类似比对程序

软件	网址	概述
BLAST	http://us.expasy.org/tools/blast/	提供在 ExPAsY 上的 BLAST 服务
BLAST	http://www.ch.embnet.org/software/BottomBLAST.html	提供在 EMBnet-CH/SIB 上的服务
BLAST	http://www.ncbi.nlm.nih.gov/BLAST/	NCBI 上的服务器
WU-BLAST	http://dove.bork.embl-heidelberg.de/Blast2/	提供在 Bork's group in EMBL
WU-BLAST	http://www.ebi.ac.uk/blast2/	提供在 EBI
BLAST	http://www.ebi.ac.uk/blastall/	提供在 EBI
Bic ultra	http://genome.dkfz-heidelberg.de/genweb/	用 SMITH-WATERMAN 算法进行相似性寻找
MPsrch	http://www.ebi.ac.uk/MPsrch/	EBI 提供的用 SMITH-WATERMAN 算法进行序列相似性比较
DeCypher	http://www.timelogic.com/decypher-interface.html	EBI 提供的用 SMITH-WATERMAN 算法进行序列相似性比较
Fasta3	http://www.ebi.ac.uk/fasta33/	EBI 提供的进行 FASTA 序列比对
FDF	http://www.ch.embnet.org/software/FDF-form.html	EMBnet-CH 上的快速寻找
SCANPS	http://www.ebi.ac.uk/scanps/	用 BATON 算法寻找相似性序列
PropSearch	http://www.embl-heidelberg.de/prs.html	寻找结构同源性

除了上述 BLAST,序列相似性比较类似的工具还有序列比对。序列比对通常指用户自定义序列的比对,从而了解几条序列之间的相似性以及它们的遗传距离。主要有两种方式:两两序列比对和多重序列连配。两两序列比对指的是两条序列之间的比较和分析,用于分析两序列之间的相似程度。多重序列连配除了能分析多条序列之间的相似程度外,还允许探察残基的变化对于蛋白质功能的影响,是常用而且重要的手段。序列比对的同时,常常需要一些图形化的软件来分析比对后的结果,如画出多重序列进化树。应用这些软件有助于了解序列间的相似程度和功能相似性。两两序列比对的工具如表 22.7 所示。

表 22.7　两两序列比对的工具

软件	网址	概述
SIM + LAL-NVIEW	http://us.expasy.org/tools/sim-prot.html	两两序列比对,并用 LAL-NVIEW 以图形的方式表示结果
Dotlet	http://www.isrec.isb-sib.ch/java/dotlet/Dotlet.html	用矩阵的方法来进行序列比较
LALIGN	http://www.ch.embnet.org/software/LALIGN-form.html	在两段序列中找出多重匹配的亚结构

多重序列连配的工具如表 22.8 所示:

表 22.8　多重序列连配的工具

软件	网址	概述
CLUSTALW	http://www.ebi.ac.uk/clustalw/	进行多重序列比对，给出相似性和遗传距离
T-Coffee	http://www.ch.embnet.org/software/TCoffee.html	多重序列比对
ALIGN	http://www2.igh.cnrs.fr/bin/align-guess.cgi	Genestream 上的多重序列比对的程序
DIALIGN	http://bibiserv.techfak.uni-bielefeld.de/dialign/	以两两序列比较为基础的多重序列比对
Match-Box	http://www.fundp.ac.be/sciences/biologie/bms/matchbox-submit.html	位于 University of Namur, Belgium
MSA	http://www.ibc.wustl.edu/ibc/msa.html	Washington University
Multalin	http://www.toulouse.inra.fr/multalin.html	多重序列比对
MUSCA	http://www.research.ibm.com/bioinformatics/musca.html	通过 patter 的寻找和比对来进行多重序列比对
AMAS	http://barton.ebi.ac.uk/servers/amas-server.html	多重序列比对
Bork's alignment tools	http://www.bork.embl-heidelberg.de:8080/Alignment/	加速多重序列比对的各种工具
CINEMA	http://www.biochem.ucl.ac.uk/bsm/dbbrowser/CINEMA2.1/	多重序列比对的图形化工具
ESPript	http://prodes.toulouse.inra.fr/ESPript/cgi-bin/nph-ESPript-exe.cgi	打印多重序列比对结果的工具
plogo	http://www.cbs.dtu.dk/~gorodkin/appl/plogo.html	位于 CBS/Denmark
WebLogo	http://www.bio.cam.ac.uk/cgi-bin/seqlogo/logo.cgi	位于 Cambridge/UK
GENIO/logo	http://genio.informatik.uni-stuttgart.de/GENIO/logo/logo.cgi	位于 Stuttgart/Germany

（五）蛋白质序列的 Pattern 和 Profiles 分析

对于分析和预测蛋白质的功能来说，简单的氨基酸组成及其比对是远远不够的。因为实际上发挥作用的是蛋白质功能域（domain，pattern，profiles，motif）。那么更进一步的工作便是分析蛋白质的功能域。现在的主流方向是通过寻找和鉴定蛋白质的功能域来分析和预测其功能以及进行分类工作。蛋白质特征数据库已经成为分析新序列的重要工具，它们主要用于蛋白质序列的分类以及推断其功能。这些亚序列的寻找和鉴定使预测的准确度大大提高。有很多关于蛋白质功能域的数据库，比如 Prosite，Smart，Prints 等等。下面简单介绍一下 InterPro[54]。它是一种整合性的数据库，包括与蛋白质功能域密切相关的 7 大数据库：PRINTS[55] 33.0 (1 650 fingerprints)，PROSITE[56] 17.25 (1587 and 162 preliminary profiles)，ProDom[57] 2001.3 (1 346 domains)，SMART[58] 3.4 (654 domains)，TIGRFAMs[59]2.1 (1614 entries)，Swiss-Prot 40.33 (118 357 entries)，TrEMBL 22.2 (732 596 entries)，Pfam[60] 8 (5 193 entries)。

目前包含了 6 725 个实体，1 453 个蛋白质功能结构域，5 121 个蛋白质家族，136 种重复序列，15 种翻译后修饰位点。这种整合使得搜索蛋白质功能域的工作变得相对完善和便利。InterPro Scan 是一种基于 Interpro 的扫描工具，它给单个用户分析新的序列，对

于从基因组计划中预测出来的蛋白质自动功能注释有很大的作用。下面介绍其具体用法。

进入 http://www.ebi.ac.uk/interpro/scan.html，在该网页的文本框中填入蛋白质序列(fasta 格式)。也可以用填写序列的文件路径方式来代替。Query Mode 选择通过 E-mail 方式或是网页交互方式得到结果。如果序列比较长或序列比较多，建议选择 Email 方式得到结果。Email 文本框中填入 Email 地址，然后点击 Submit，递交后弹出一小窗口，显示运行的时间。完成后窗口会自动关闭，同时显示出结果，如图 22.5。



图 22.5　Interproscan 界面

图 22.5 可见，通过 Interproscan 找到了与此蛋白质相似的蛋白质家族信息(PR01553)和功能结构域信息(PF03507)。点击中间 IPR 开头的链接，给出的是 Interpro 数据库中的实体信息。

每一个 InterPro 实体有以下几种组成。

① accession number：单一序列号。描述和这个实体相联系的蛋白质特征和文字参考，并和相关数据库联系起来。它的格式是 IPRXXXXXX。

② name：命名。简单描述相匹配的实体。

③ type：类型。表示 family，domain，repeat，或者 PTM(转录后修饰)中某一种类型。

④ signatures：列出了相关数据库的超级链接。

⑤ function：分子功能。简单描述这个蛋白质的功能，并给出 GO 序列号。

⑥ abstract：提供了相关功能的文字参考。

⑦ example：给出了位于同一个实体下蛋白质的例子。

另外，点击 P 开头的超级链接到达所对应的网站，比如点击 PF03507，则到达 PFAM 查询网站，显示相应的 PFAM 的结构信息。

得到了相应蛋白质功能域的信息以后，如果你还觉得不够详细，可以通过 InterPro 链接到 GO 数据库。GO 全称是：Gene Ontology(http://www.geneontology.org)，它是对目前存在的基因产物进行分析和定位的数据库，能从功能、组成、细胞过程三个方面来分析蛋白质。可以到相应的网站参考其帮助文档。软件及数据库如表 22.9 所示。

表 22.9 蛋白质功能域分析工具

软件	网址	概述
Interpro Scan	http://www.ebi.ac.uk/interpro/scan.html	用于扫描 PROSITE, PRINTS 等多种蛋白质家族的数据库的整合工具
ScanProsite	http://us.expasy.org/tools/scanprosite/	扫描 PROSITE 或者是 Swis-sProt, TrEMBL 的工具
MotifScan	http://hits.isb-sib.ch/cgi-bin/PFSCAN?	在蛋白质家族数据库中扫描一段序列
Frame-ProfileScan	http://www.isrec.isb-sib.ch/software/PFRAMESCAN-form.html	与 MotifScan 类似,只是扫描的序列较短
Pfam HMM search	http://www.sanger.ac.uk/Software/Pfam/search.shtml	在 Pfam 数据库中扫描和寻找序列
finger-PRINTScan	http://www.bioinf.man.ac.uk/fingerPRINTScan/	在 PRINTS 数据库中扫描和寻找序列
FPAT	http://stateslab.bioinformatics.med.umich.edu/service/fpat/	在蛋白质数据库中寻找规律表达的序列
PRATT	http://www.ebi.ac.uk/pratt/	在一些未进行比对的序列中找出具有保守特征的区域
PPSEARCH	http://www.ebi.ac.uk/ppsearch/	扫描 PROSITE 数据库.(EBI)
PROSITE SCAN	http://npsa-pbil.ibcp.fr/cgi-bin/npsa-automat.pl? page = npsa-prosite.html	扫描 PROSITE 数据库.(PBIL)
PATTERN-PROT	http://npsa-pbil.ibcp.fr/cgi-bin/npsa-automat.pl? page = npsa-pattinprot.html	扫描蛋白质数据库以寻找一个或者几个 pattern.(PBIL)
SMART	http://smart.embl-heidelberg.de/	简单分子结构研究工具.数据库.(EMBL)
TEIRESIAS	http://www.research.ibm.com/bioinformatics/teiresias.html	从一系列未比对的蛋白质或者核酸中找出 pattern
Hits	http://hits.isb-sib.ch/	蛋白质序列和 motifs 之间的关系分析

(六) 亚细胞定位

亚细胞定位分析对于预测蛋白质功能的作用不可小觑。它不仅有助于深入研究蛋白质的功能,对于整个基因组功能分析也不无裨益,还可以了解和研究各个细胞器的进化过程。如果得到一段蛋白质序列,可以用 SignalP 和 TargetP、PSort 来预测蛋白质定位。

下面以 Signal P[61~67]和 TargetP[61~68]为例详细说明软件的用法和结果分析。

1. Signal P 软件

Signal P 是专门用于信号肽预测的软件。它用人工神经网络的算法来预测不同种类蛋白质序列中的信号肽及其剪切位点。分为三种:革兰氏阳性原核生物、革兰氏阴性原核生物和真核生物。

目前主要是在网络上提交氨基酸序列的形式来运行。输入序列后,选择生物种类,程序会反馈给你三个在 0 与 1 之间的分数:

C-SCORE：剪切位点分数。分数越高，表示是剪切位点的可能性越大。

S-SCORE：信号肽分数。分数低则表示所输入的这段序列是在剪切位点的后 30 个氨基酸之后，或是非分泌蛋白质的 N 端。

Y-SCORE：联合分数。这是将 S-SCORE 和 C-SCORE 联合分析后得出的最优结果。对于每一条序列来说，Signal P 都会给出 C 和 Y 分数的最大值，以及 N 端和预测的剪切位点之间的平均 S 分数。这些值用于区别信号肽和非信号肽。如果序列被预测有信号肽，那么剪切位点应该在具有最大 Y 值的位点前面。图 22.6～图 22.8 为详细的分析结果。

① 对于一个典型的信号肽来说，C 值和 Y 值应该在 1 的位置，同时平均 S 值在剪切位点之前都很高，之后突然变低。如图 22.6。

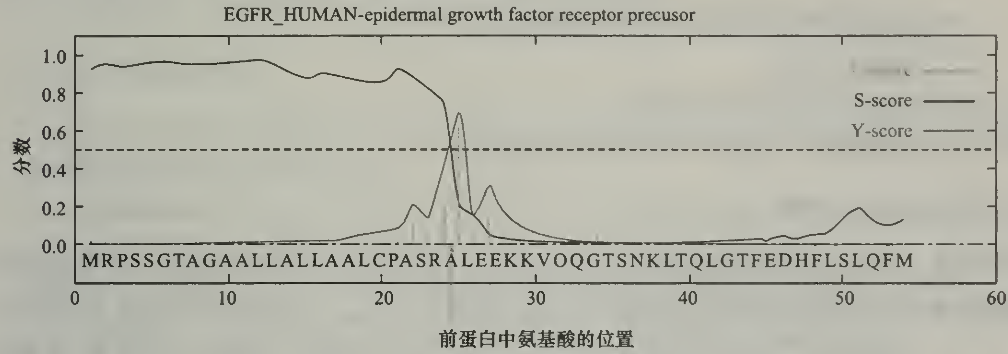


图 22.6　Signal P 分析结果 1

② 图 22.7 表示的例子中，如果只看 C 值的最高点，真实的剪切位点会被错误的预测。但是 Signal P 联合 Y 值还是可以正确预测。

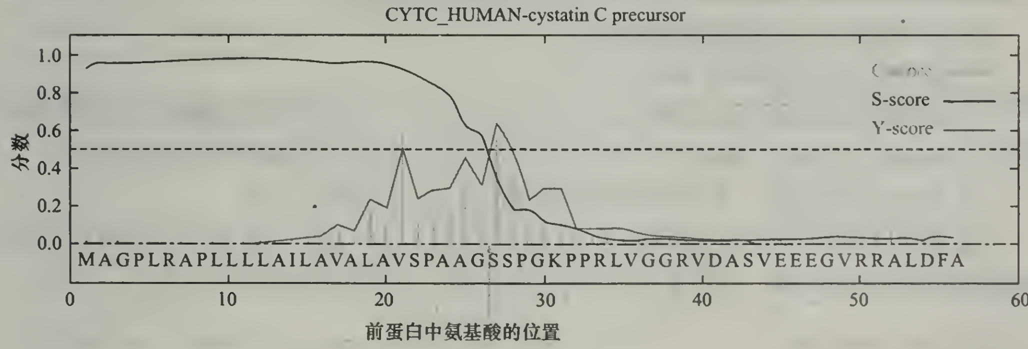


图 22.7　Signal P 分析结果 2

③ 对于典型的非分泌蛋白质来说，所有的三个分数都很低。如图 22. 8。

以上介绍的是 SignalP 1.1 版。现在 SignalP 已经更新到 2.0 版。相对于旧版本来说，2.0 版有了很多改进。主要有以下几个方面。

① 结合了隐马尔可夫模型(HMM)算法，预测的准确率有所提高。

② 用 Swiss-Prot Rel.35 中的数据重新训练了神经网络(相对于 1.1 版的 Swiss-Prot Rel 29)。

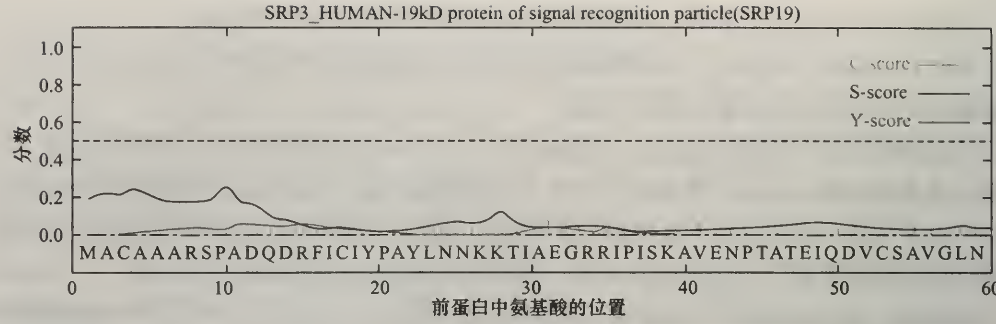


图 22.8　Signal P 分析结果 3

③ 结果输出整合了简单、方便的图形方式。

④ 自动修订数据。

⑤ 信号肽区域的详细功能分析。

⑥ 可以将程序本土化，1.1 版只有网络版本。

2. Target P 软件

Target P 软件用来预测蛋白质的亚细胞定位以及潜在信号肽剪切位点。这些定位包括线粒体、叶绿体、内质网/高尔基体分泌蛋白质等。Target P 注意事项和用法如下所述。

① 一般情况下，提交的序列都要含有 N 端。离 N 端越远，预测的难度越高，准确性越低。

② 提交序列的推荐长度是 130 或以上 N 端残基，因为在 Target P 训练集中用的是 130 个或以上的残基长度。如果在 100 或 120 以内，cTP(chloroplast transit peptide)，mTP (mitochondria transit peptide) 预测的准确度就要受到影响。

③ 序列命名可以在任何长度范围内。

④ 提交序列格式必须是 FASTA 格式。

⑤ 根据所要预测的蛋白质来源选择预测的内容。非植物种类预测内容包括线粒体定位，分泌途径及其他。除此以外，植物种类预测内容还包括叶绿体的可能性定位。

⑥ 剪切位点的选择。如果你想预测剪切位点，可以选择这一项。

⑦ 最后是选择 cutoff 值，在默认的状态下，没有 cutoff 值。如果选择了特异性，相应会给出设定的 cutoff 值。还可以自己设定 cutoff 值。执行结果如图 22.9 所示。

TargetP V1.0 Prediction Results

```
### ### ###   T A R G E T P   1.0   prediction results   ### ### ###

# Number of input sequences: 1
# Cleavage site predictions included.
# Using NON-PLANT networks.

#                        Name    Length    mTP   SP  other  Loc.  Rc    TPlen
#----------------------------------------------------------------------------
                       human_1       245 0.068 0.107 0.926   _     1     -
#----------------------------------------------------------------------------
# cutoff                                  0.00  0.00  0.00
```

图 22.9　Target P 分析结果

图 22.9 给出了提交序列的名字和长度，以及每一个可能性定位的分数。真正的预测定位是在 loc 这一列给出的。分别用 5 个字母来表示：C：叶绿体；M：线粒体；S：分泌途径；-：除了以上三种以外的任何其他的定位；*：未知。

要注意到三种定位分数并不是真正的预测结果，而且没有必要相加等于 1。分数最高的便是可能性最大的定位。

RC 为可信度。其衡量标准是最高分和第二高分之间的差值。差值越大，RC 数值越低，表示可信度越高。级别设置如下：

RC＝1：差值＞0.800；RC＝2：0.800＞差值＞0.600；RC＝3：0.600＞差值＞0.400；RC＝4：0.400＞差值＞0.200；RC＝5：差值＜0.200。

如果选择了预测剪切位点，TPlen 表示的就是所要剪切序列的长度。

3. 其他类似软件

Chloro P[68]用于预测定位于叶绿体中的蛋白质。MitoProt[69]用于预测定位于线粒体中的蛋白质（EBI 的分析软件）。

上述软件中，Chloro P、MitoProt、Signal P 对于其特定的蛋白质定位，敏感性比 TargetP 高，但是相应的特异性却比较低。Target P 对于剪切位点的预测不够好，准确率只能达到 40%～50%。所以，如果需要预测确切的剪切位点，应该选用特定的软件。Psort 同样可以实现上述功能，也是常用的软件。亚细胞定位软件总结如表 22.10。

表 22.10 亚细胞定位软件

软件	网址	概述
SignalP	http://www.cbs.dtu.dk/services/SignalP/	预测信号肽及其剪切位点
ChloroP	http://www.cbs.dtu.dk/services/ChloroP/	预测转运到叶绿体的肽段
MitoProt	http://www.mips.biochem.mpg.de/cgi-bin/proj/medgen/mitofilter	预测转运到线粒体的肽段
Predotar	http://www.inra.fr/predotar/	预测定位到线粒体和胞浆中的肽段
PSORT	http://psort.ims.u-tokyo.ac.jp/	预测蛋白质分选信号和亚细胞定位的位点
TargetP	http://www.cbs.dtu.dk/services/TargetP/	预测蛋白质亚细胞定位位点
DAS	http://www.sbc.su.se/~miklos/DAS/	预测原核细胞蛋白质中的跨膜区域
HUMMTOP	http://www.enzim.hu/hmmtop/	预测蛋白质的跨膜螺旋和拓扑结构
PredictProtein	http://cubic.bioc.columbia.edu/predictprotein/	预测蛋白质的跨膜螺旋和拓扑结构
SOSUI	http://sosui.proteome.bio.tuat.ac.jp/sosuiframe0.html	预测跨膜区域
TMAP	http://www.mbb.ki.se/tmap/	通过多序列连配预测蛋白质跨膜区域
TMHMM	http://www.cbs.dtu.dk/services/TMHMM-2.0/	预测蛋白质中的跨膜螺旋
TMPRED	http://www.ch.embnet.org/software/TMPRED-form.htm	预测跨膜螺旋和蛋白质定位
TOPPRED	http://www.sbc.su.se/~erikw/toppred2/	预测膜蛋白质的拓扑结构

（七）蛋白质相互作用(PPI)

预测蛋白质相互作用的软件正处于发展中。作为研究蛋白质相互作用的软件 DIP 功能比较全面[72~77](http://dip.doe-mbi.ucla.edu/)。它是一个描述已知蛋白质相互作用的数据库,具体是指实验上已经确认了的可以相互结合和作用的氨基酸序列。该数据库建立的目的是为了研究特殊的蛋白质相互作用,以及研究探索整个调节和信号转导途径,还可以从整个细胞水平上来研究蛋白质相互作用网络的整体性和复杂性。想要了解具体情况可以参考 http://dip.doe-mbi.ucla.edu/上的帮助文档。蛋白质相互作用分析软件如表 22.11 表示。

表 22.11　蛋白质相互作用分析软件

软件	网址	概述
JDIP2D	http://dip.doe-mbi.ucla.edu/dip/jdip.cgi	除了提供一个 DIP(database of interacting proteins)的接口以外,还能对于蛋白质相互作用的网络进行相应的注释
LiveDIP	http://dip.doe-mbi.ucla.edu/ldipc/tmpl/livedip.cgi	是 DIP 的扩展,主要运用蛋白质的状态及其中间态来描述逻辑上的蛋白质相互作用
DLRP	http://dip.doe-mbi.ucla.edu/dip/DLRP.cgi	是 DIP 的子集,主要是关于配体和受体相互作用的数据库
Protein-Protein Interaction server	http://www.biochem.ucl.ac.uk/bsm/PP/server	预测蛋白质蛋白质相互作用

（八）二级结构预测

目前,预测蛋白质二级结构的技术已比较成熟。二级结构预测方法大致可以分为两类:一类是根据单一序列来预测二级结构,另一类是根据多序列列线来预测二级结构。前者如著名的 GOR 算法,也只有 60%左右的预测准确率。特别是对于富含 β 折叠的序列,准确率更低。主要原因是该算法忽略了氨基酸残基的空间结构排列,而实际上这种空间排列对于蛋白质的功能是非常重要的。而多序列列线则改变了这一现状,它把空间排列作为考虑的重点,使得预测的准确率提高了 10%左右。二级结构预测软件如表 22.12。

表 22.12　蛋白质二级结构预测软件

软件	网址	概述
PredictProtein	http://cubic.bioc.columbia.edu/predictprotein/	预测二级结构,在 Columbia University
BCM PSSP	http://dot.imgen.bcm.tmc.edu:9331/pssprediction/pssp.html	Baylor College of Medicine
Prof	http://www.aber.ac.uk/~phiwww/prof/	二级结构预测的多重分类器
HNN	http://npsa-pbil.ibcp.fr/cgi-bin/npsa-automat.pl? page = npsa-nn.html	用神经网络的分析方法预测二级结构
Jpred	http://www.compbio.dundee.ac.uk/~www-jpred/	University of Dundee 中运用整合的方法来预测的软件
nnPredict	http://www.cmpharm.ucsf.edu/~nomi/nnpredict.html	University of California at San Francisco
PSA	http://bmerc-www.bu.edu/psa/request.htm	BioMolecular Engineering Research Center
PSIpred	http://bioinf.cs.ucl.ac.uk/psipred/	多种预测方法 Brunel University

（九）三级结构的预测

预测三维结构有三种手段：一是 *Ab-Initio*。由于计算手段的限制，准确率一直达不到预测结构的要求，发展的空间有限。第二种是通过折叠方式的识别来预测。第三种方法是通过结构同源性的比对来实现的。通常的做法就是使用 BLASTP 程序比对蛋白质 3D 数据库（比如 NRL-3D），如果发现有 40% 以上相似的匹配序列，则认为两者之间有非常相似的结构，同时可以进行同源性建模。Swiss-Model[77~88]就是建模工具中的一个代表。三级结构的预测软件如表 22.13。

表 22.13 蛋白质三级结构的预测软件

软件	网址	概述
SWISS-MODEL	http://us.expasy.org/swissmod/SWISS-MODEL.html	自动建模工具
Geno3d	http://geno3d-pbil.ibcp.fr/	自动三维结构建模工具
CPHmodels	http://www.cbs.dtu.dk/services/CPHmodels/	以神经网络为基础的自动建模工具
3D-PSSM	http://www.sbg.bio.ic.ac.uk/~3dpssm/	联合一级、二级、三级结构来识别蛋白质折叠的信息
SWEET	http://www.dkfz-heidelberg.de/spec/sweet2/doc/index.html	从序列信息来进行三维结构的建模
Swiss-Pdb-Viewer	http://us.expasy.org/spdbv/	展示蛋白质三维结构的工具
RasMol	http://www.umass.edu/microbio/rasmol/	展示三维结构

（十）其　　他

在生物信息学软件中，有很多小的软件能帮助用户解决很多手边的小问题。如果能了解得多一些，可以明显提高工作效率。比如计算分子质量、转换氨基酸表达形式，等等。另外，一些结构可视化的软件可以将复杂的生物学数据转化为简单的图形，有助于全面、系统地分析。其他工具见表 22.14。

表 22.14 其他工具

软件	网址	概述
Drawhca	http://smi.snv.jussieu.fr/hca/hca-form.html	画出蛋白质序列中的疏水区域
Protein Colourer	http://www.ebi.ac.uk/cgi-bin/translate/visprot.pl	给氨基酸序列涂上不同的颜色
Three to one	http://www.ualberta.ca/~stothard/javascript/threeToOne.html	将三个字母表示的氨基酸序列转换为一个字母表示
Colorseq	http://npsa-pbil.ibcp.fr/cgi-bin/npsa-automat.pl?page=npsa-color.html	突出所选择氨基酸残基
RandSeq	http://us.expasy.org/tools/randseq.html	随意造出一段蛋白质序列
Helix Wheel	http://www.site.uottawa.ca/~turcotte/resources/HelixWheel/	用螺旋的方式来表示蛋白质序列

四、实验室信息管理系统

蛋白质组学研究高通量这一个特性需要一整套完整的规范来进行数据的采集、整理、分析。利用计算机技术可以建立实验室信息管理系统，通过计算机管理可以使实验规范化，数据的采集管理和分析高度自动化。不同的实验室有不同的需求，具体表现为不同的实验方法、不同的实验流程、不同的软件处理以及不同的数据标准。尽管各个实验室的信息管理系统看上去千差万别，采用的架构方式也不尽相同，但实质上，实验室信息管理系统主要分两个部分：实验数据管理采集和数据分析。实验数据管理采集部分主要应该参考计算机工程中比较成熟的 MIS（管理信息系统），其核心思想是尽量减少人的干预，将人的工作规范化，信息处理的自动化。数据的分析部分主要是对实验中采集到数据作鉴定和数据挖掘等工作，这是最具各个实验室特色的部分，不同的实验室有不同的数据要求。实验室信息管理系统的系统架构大同小异，以系统架构来分主要由 BS（浏览器-服务器）架构或者是 CS（客户端-服务器）架构，这两种架构都是由用户交互界面加数据库完成的。以数据的流程来讲可以分为：数据的输入，数据的存储与管理，有选择的提取数据，对提取的数据进行分析，分析结果的显示和再入库。

我们实验室也建立了自己的实验室信息管理系统（Labmis）。实验室信息管理系统涉及到实验室工作的方方面面，如何有效地进行组织和管理不是单纯的计算机问题或生物学问题，它更多体现了一个实验室的管理水平，需要生物学家对实验室的方方面面进行系统的分析，妥善的规划，才能使得开发出来的实验室信息管理系统得到最有效的利用，否则将只是一个摆设（还要占用时间和空间）。一个高效的采用计算机工程技术进行系统分析设计的方法就是采用 UML（统一建模语言）来建模。UML 是计算机设计建模的工具之一，与生物信息学关系不大，可以参考相关计算机书籍进行学习。

由于各个实验室的数据存放的格式方式，采用的标准都不尽相同，给研究者在数据的交换和共享方面带来一定的困难。为了便于科学家们进行数据交换和共享，国际蛋白质组组织 HUPO 组织了全世界研究蛋白质组的科学家来设计 2D 数据交换和数据库存储的格式。数据格式本身采用计算机技术中成熟的可扩展标记语言 XML（XML 类似 HTML 的标记语言，但和 HTML 不同的是，根据 XML 的定义，它不限于一个特定的词汇表、行业或者用途。XML 可以很方便地在数据库之间进行数据的转移工作）。XML 既可以作为数据存储的载体又可作为数据交换的载体，而且通用数据库基本都支持 XML。XML 蛋白质组数据格式的设计满足两个基本原则：一是满足公开发表的最小需求，二是满足数据比较的最小需求。整个数据交换格式的设计思路是：目的需求、样品信息、样品制备、实验方法、实验结果、蛋白质斑点信息，等等。

XML 数据交换格式的组成分为 3 个部分：第一部分是样品和双向凝胶电泳实验部分，主要有样品信息、2D 胶信息、1D 胶信息、蛋白质斑点信息等表组成。第二部分为蛋白质鉴定实验部分，包含了质谱类型，检测方法等表结构。第三部分是蛋白质检测的信息，包含了蛋白质斑点最主要的质谱检测的结果（肽段的核质比）的信息。

使用 XML 数据交换格式的另一个好处是：可以编写不同的样式表单或各式文档，非常方便地通过普通浏览器或者 XML 专用浏览器显示数据。当然也可以专门写一个

XML 的应用程序来显示和修改 XML 数据。

五、总　　结

从某种意义而言,生命系统就是一个庞大的信息处理、传递和整合系统,它的复杂程度远远超过任何人工设计的系统。传统的生命科学研究在采用个体和手工方式研究简单系统时,产生的数据规模是小型的,对处理海量数据的需求不大,认识不深。但我们必须意识到,从基因组计划一开始启动,生命科学的研究就进入到大科学、系统生物学的时代,各种大通量、系统性的研究手段将生命信息以前所未有的规模和数量呈现在我们面前,这才是生命存在的真实状态。传统的低通量、人依赖性的方法显然不再适用,取而代之的是高通量、机器依赖的研究方式,这就是生物信息学。一方面,生物信息学作为一门新兴的学科,本身还需要发展和积累,另一方面,我们面临着各种大科学不断兴起,新的问题不断产生的时代,对生物信息学的需求与日俱增。基因组计划的开展已经大大推动了生物信息学的发展,但蛋白质组学研究比基因组更复杂并提出了许多新问题,如蛋白质-蛋白质相互作用、蛋白质定位等。这些问题的解决,都不同程度地依赖着生物信息学的发展,因此,可以认为,蛋白质组生物信息学研究还处于起步阶段,面临着大量挑战,拥有着广阔的发展前景。

（陈益强　付　岩　李亦学　李素君　曾　嵘　张　雷）

参考文献

1 Perkins DN, Pappin DJ, Creasy DM *et al*. Probability-based protein identification by searching sequence databases using mass spectrometry data. Electrophoresis, 1999, 20: 3551～3567

2 Eng JK, McCormack AL, Yates JR. An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. J Am Soc Mass Spec, 1994, 5: 976

3 Field HI, Fenyö D, Beavis RC. RADARS, a bioinformatics solution that automates proteome mass spectral analysis, optimises protein identification, and archives data in a relational database. Proteomics, 2002, 2(1): 36～47

4 Bafnav, Edwards N. SCOPE, a probabilistic model for scoring tandem mass spectra against a peptide database. Bioinformatics, 2001, 17(Supp. 1): 13～21

5 Clauser KR, Baker P, Burlingame AL. Role of accurate mass measurement (±10 ppm) in protein identification strategies employing MS or MS/MS and database searching. Anal Chem, 1999, 71: 2871～2882

6 Mann M, Wilm M. Error～tolerant identification of peptides in sequence databases by peptide sequence tags. Anal Chem, 1994, 66: 4390～4399

7 Fenyö D, Qin J, Chait BT. Protein identification using mass spectrometric information. Electrophoresis, 1998, 19: 998～1005

8 Cottrell JC, Green BN, MaxEnt. an essential maximum entropy based tool for interpreting multiply-charged electrospray data. Micromass UK Limited. Available at http://www.micromass.co.uk/

9 Baginsky S, Cieliebak M, Gruissem W *et al*. A tool for automatic de novo peptide sequencing. Tech Reports, 2002, 10: 383

10 Yates J, Eng JK, McCormack AL *et al*. Method to correlate tandem mass spectra of modified peptides to amino acid sequences in the protein database. Anal Chem, 1995, 67: 1426～1436

11 Dancik V, Addona TA, Clauser KR *et al*. De novo peptide sequencing via tandem mass Spectrometry: a graph-theoretical

approach, 1999, RECOMB 99: 135～144
12 Taylor JA, Johnson RS. Implementation and uses of automated de novo peptide sequencing by tandem mass spectrometry. Anal Chem, 2001.73:2594～2604
13 Goodlett DR, Keller A, Watts JD *et al*. Differential stable isotope labeling of peptides for quantitation and de novo sequence derivation. Rapid Commun Mass Spectrom, 2001, 15:1214～1221
14 Fernandez-de-Cossio J, Gonzalez J, Takao T *et al*. A software program for the rapid sequence analysis of unknown peptides involving modifications, Based on MS/MS Data. 45th ASMS Conference on Mass Spectrometry and Allied Topics, Slot 074 (on-line proceedings at www.asms.org/confASMS.php), 1997
15 Taylor JA, Johnson RS. Sequence database searches via de novo peptide sequencing by tandem mass spectrometry. Rapid Comm Mass Spec, 1997, 11:1067～1075
16 http://www.immunex.com/researcher/lutefisk/
17 Baginsky S, Cieliebak M, Gruissem W *et al*. A tool for automatic de novo peptide sequencing. Technical Report no, 383, ETH Zurich, Dept. of Computer Science
18 http://www.appliedbiosystems.com
19 http://www.daltonics.bruker.com
20 http://www.thermo.com
21 http://www-hto.usc.edu/ tingchen/resume.html
22 http://www.thermo.com/eThermo/CDA/Products/Product-Detail/1,1075,19050～113,00.html
23 Chen T, Kao M Y, Tepel M *et al*. A dynamic programming approach to de novo peptide sequencing via tandem mass spectrometry Proceeding of the 11th SIAM-ACM Symposium on Discrete Algorithms (SODA 2000), 2000: 389
24 Ma B, Zhang K, Lajoie G *et al*. A powerful software tool for the de novo sequencing of peptides from MS/MS data. Abstracts of the 50th ASMS Conference on Mass Spectrometry and Allied Topics, 2002, 2(6): 70
25 http://www.bioinformaticssolutions.com
26 Perkins DN, Pappin DJC, Creasy D *et al*. Probability based protein identification by searching sequence databases using mass spectrometry data. Electrophoresis, 1999, 20: 3551～3567
27 http://fields.scripps.edu/sequest/
28 Bartels C. Fast algorithm for peptide sequencing by mass spectroscopy biomed. Environ Mass Spectrom, 1990, 19: 363～368
29 Dancik V, Addona T, Clauser K *et al*. de novo peptide sequencing via tandem mass spectrometry: a graph-theoretical approach. J Comput Biol, 1999, 6:327～341
30 Chen T, Kao M, Tepel M *et al*. A dynamic programming approach to de novo peptide sequencing via tandem mass spectrometry. J Comput Biol, 2001, 8(3):325～337
31 Rogers S, Wells R, Rechsteiner M. Amino acid sequences common to rapidly degraded proteins: the PEST hypothesis. Science, 1986, 234:364～368
32 Rechsteiner M, Rogers SW. PEST sequences and regulation by proteolysis. TIBS, 1996, 21:267～271
33 Wilkins MR, Gasteiger E, Gooley A *et al*. High-throughput mass spectrometric discovery of protein post-translational modifications. J Mol Biol, 1999, 289:645～657
34 Altschul SF, Gish W, Miller W *et al*. Basic local alignment search tool. J Mol Biol, 1990, 215:403～410
35 Gish W, States DJ. Identification of protein coding regions by database similarity search. Nat Genet, 1993, 3:266～272
36 Madden TL, Tatusov RL, Zhang J. Applications of network BLAST server. Meth Enzymol, 1996, 266:131～141
37 Altschul SF, Madden TL, Schäffer AA *et al*. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res, 1997, 25:3389～3402
38 Zhang J, Madden TL. PowerBLAST: A new network BLAST application for interactive or automated sequence analysis and annotation. Genome Res, 1997, 7:649～656
39 Altschul SF, Gish W. Local alignment statistics. Meth Enzymol, 1996, 266:460～480
40 Wootton JC, Federhen S. Analysis of compositionally biased regions in sequence databases. Meth Enzymol, 1996, 266:

554～571

41 Altschul SF, Boguski MS, Gish W *et al*. Issues in searching molecular sequence databases. Nat Genet, 1994, 6: 119～129

42 Wootton JC, Federhen S. Statistics of local complexity in amino acid sequences and sequence databases. Comput Chem, 1993, 17: 149～163

43 Hancock JM, Armstrong JS. SIMPLE34: an improved and enhanced implementation for VAX and Sun computers of the SIMPLE algorithm for analysis of clustered repetitive motifs in nucleotide sequences. Comput Appl Biosci, 1994, 10: 67～70

44 Dayhoff MO, Schwartz RM, Orcutt BC. A model of evolutionary change in proteins. Atlas of Protein Sequence and Structure, 1978, 5: 3supply

45 Schwartz RM, Dayhoff MO. Matrices for detecting distant relationships. Atlas of Protein Sequence and Structure, 1978, 5: suppl. 3

46 Altschul SF. Amino acid substitution matrices from an information theoretic perspective. J Mol Biol, 1991, 219: 555～565

47 States DJ, Gish W, Altschul SF. Improved sensitivity of nucleic acid database searches using application-specific scoring matrices. Methods, 1991, 3: 66～70

48 Henikoff JG. Amino acid substitution matrices from protein blocks. Proc Natl Acad Sci USA, 1992, 89: 10915～10919

49 Altschul SF. A protein alignment scoring system sensitive at all evolutionary distances. J Mol Evol, 1993, 36: 290～300

50 Karlin S, Altschul SF. Methods for assessing the statistical significance of molecular sequence features by using general scoring schemes. Proc Natl Acad Sci USA, 1990, 87: 2264～2268

51 Karlin S, Altschul SF. Applications and statistics for multiple high-scoring segments in molecular sequences. Proc Natl Acad Sci USA, 1993, 90: 5873～5877

52 Dembo A, Karlin S, Zeitouni O. Limit distribution of maximal non-aligned two-sequence segmental score. Ann Prob, 1994, 22: 2022～2039

53 Altschul SF. Evaluating the statistical significance of multiple distinct local alignments. Theoretical Comput Methods Genome Res, 1997, 23: 1～14

54 Mulder NJ, Apweiler R, Attwood TK *et al*. The interPro database, 2003 brings in creased coverage and new features. NAR, 2003, 31: 315～318

55 Attwood TK, Bradley P, Flower DR *et al*. PRINTS and its automatic supplement, prePRINTS. NAR, 2003, 31: 400～402

56 Falquet L, Pagni M, Bucher P *et al*. The PROSITE database, its status in 2002. NAR, 2002, 30: 235～238

57 Corpet F, Servant F, Gouzy J *et al*. ProDom and ProDom-CG: tools for protein domain analysis and whole genomecomparisons, NAR, 2000, 28: 267～269

58 Letunic I, Goodstadt L, Dickens NJ *et al*. Recent improvements to the SMART domain-based sequence annotation resource. NAR, 2002, 30: 242～244

59 Haft DH, Selengut JD, White O. The TIGRFAMs database of protein families. NAR, 2003, 31: 371～373

60 Bateman A, Birney E, Cerruti L *et al*. The Pfam protein families database. NAR, 2002, 30: 276～280

61 Henrik N, Jacob E, Søren B *et al*. Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. Pro Engineer, 1997, 10: 1～6

62 Henrik N, Anders K. Prediction of signal peptides and signal anchors by a hidden Markov model. Proceedings of the Sixth International Conference on Intelligent Systems for Molecular Biology (ISMB 6). AAAI Press, Menlo Park, California, 1998, 122～130

63 Henrik N, Søren B, Gunnar H. Machine learning approaches to the prediction of signal peptides and other protein sorting signals. Prot Engin, 1999, 12: 3～9

64 Henrik N, Jacob E, Søren B *et al*. A neural network method for identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. J Neural Sys, 1997, 8: 581～599

65 Henrik N. From sequence to sorting: prediction of signal peptides. Defended at Department of Biochemistry, Stockholm University, Sweden, May 25 1999

66 Henrik N. Hot Paper in bioinformatics-an interview by Eugene Russo with SignalP author Henrik Nielsen . The Scientist, 1999, 13:8

67 Olof E, Henrik N, Søren B *et al*. Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. J Mol Biol, 2000, 300: 1005～1016

68 Olof E, Henrik N, Gunnar H. ChloroP, a neural network-based method for predicting chloroplast transit peptides and their cleavage sites. Prot Sci, 1999, 8: 978～984

69 Claros MG, Vincens P. Computational method to predict mitochondrially imported proteins and their targeting sequences. Eur J Biochem, 1996, 241: 779～786

70 Xenarios I, Salwinski L, Duan XJ *et al*. DIP: the database of interacting proteins. A research tool for studying cellular networks of protein interactions. NAR, 2002, 30: 303～305

71 Xenarios I, Fernandez E, Salwinski L *et al*. DIP: the database of interacting proteins: 2001 update. NAR, 2001, 29: 239～241

72 Xenarios I, Rice DW, Salwinski L *et al*. DIP: the database of interacting proteins. NAR, 2000, 28: 289～291

73 Deane CM, Salwinski L, Xenarios I *et al*. Protein interactions: two methods for assessment of the reliability of high-throughput observations. Mol Cell Prot, 2002, 1: 349～356

74 Graeber TG, Eisenberg D. Bioinformatic identification of potential autocrine signaling loops in cancer using gene expression profiles. Nat Genet, 2001, 29: 295～300

75 Peitsch MC. ProMod: automated knowledge-based protein modelling tool. PDB Quarterly Newsletter, 1995, 72: 4

76 Peitsch MC. Jongeneel V. A 3-dimensional model for the CD40 ligand predicts that it is a compact trimer similar to the tumor necrosis factors. Int Immunol, 1993, 5: 233～238

77 Peitsch MC. Protein modelling by E-Mail. BioTechnology, 1995, 13: 658～660

78 Peitsch MC. ProMod and Swiss-Model: Internet-based tools for automated comparative protein modelling. Biochem Soc Trans, 1996, 24: 274～279

79 Peitsch MC, Herzyk P. Molecular modelling of G-protein coupled receptors. In: G Protein-coupled Receptors. New opportunities for commercial development, 1996, 6: 29～37

80 Peitsch MC, Herzyk P, Wells TNC. Automated modelling of the transmembrane region of G-protein coupled receptor by Swiss-Model. Receptors and Channels, 1996, 4: 161～164

81 Peitsch MC, Wilkins MR, Tonella L *et al*. Large scale protein modelling and integration with the SWISS-PROT and SWISS-2DPAGE databases: the example of Escherichia coli. Electrophoresis, 1997, 18: 498～501

82 Peitsch MC. Large scale protein modelling and model repository. Proceedings of the Fifth International Conference on Intelligent Systems for Molecular Biology, 1997, 5: 234～236

83 Peitsch MC, Guex N. Large-scale comparative protein modelling. In: Proteome research: new frontiers in functional genomics, 1997, 177～186

84 Guex N, Peitsch MC. SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modelling. Electrophoresis, 1997, 18: 2714～2723

85 Guex N, Peitsch MC. Molecular modelling of proteins. Immunol News, 1999, 6: 132～134

86 Guex N, Diemand A, Peitsch MC. Protein modelling for all. TIBS, 1999, 24: 364～367

87 Peitsch MC, Schwede T, Guex N. Automated protein modelling—the proteome in 3D. Pharmacogenomics, 2000, 1: 257～266

88 Schwede T, Diemand A, Guex N *et al*. Protein structure computing in the genomic era. Res Microbiol, 2000, 151: 107～112

第二十三章　蛋白质组学在各领域的应用

第一节　蛋白质组学在微生物研究中的应用

一、概　　述

(一) 微生物简介

微生物不是生物分类系统中的一个类群,而是个体直径小于 1mm 的生物的总体,包括了至少 100 万种的生物,但它们都具有以下五大特征[1]:第一,体积小,表面积大;第二,吸收多,转化快;第三,生长旺,繁殖快;第四,适应性强,易变异;第五,分布广,物种种类多,代谢产物种类多。微生物的分类方法有很多,例如,分成细菌、真菌、病毒、藻类和原生动物五类。这里,我们依据现行微生物学教材,按照微生物的结构、化学组成及生活习性等差异将微生物分成三大类[2]。第一类,真核细胞型微生物:细胞核的分化程度较高,有核膜、核仁和染色体;胞质内有完整的细胞器(如内质网、核糖体及线粒体等)。真菌(如酵母菌、霉菌和黏菌等)、原生动物属于此类型微生物。第二类,原核细胞型微生物:细胞核分化程度低,仅有原始核质,没有核膜与核仁;细胞器不很完善。这类微生物种类众多,有细菌、支原体、立克次体、衣原体、螺旋体和放线菌等。第三类,非细胞型微生物:没有典型的细胞结构,亦无产生能量的酶系统,只能在活细胞内生长繁殖。病毒(含亚病毒)属于此类型微生物,其中亚病毒(subvirus)是指不具有完整的病毒结构的一类病毒,包括类病毒、卫星 RNA、朊病毒。

绝大多数微生物对人类和动植物的生存是有益而必需的[2]。其一,自然界中氮、碳、硫等多种元素循环靠微生物的代谢活动来进行。其二,在农业方面,人类广泛利用一些微生物的特性,开辟了以菌造肥、以菌催长、以菌防病、以菌治病等农业增产新途径。其三,在工业方面,微生物在食品、制革、纺织、石油、化工等领域的应用越来越广泛。尤其是在医药工业方面,几乎所有的抗生素都是微生物的代谢产物。另外,还可利用微生物来制造一些维生素、辅酶等药物。当然,还有一小部分微生物能引起人类或动植物的病害,这些具有致病性的微生物称为病原微生物。有些微生物在正常情况下不致病,而在特定条件下可引起疾病,称为条件性病原微生物。

由于微生物个体微小、种类繁多,人们对其认识和研究经历了漫长的岁月。直到 20 世纪 50 年代初,"DNA 双螺旋结构模型"的建立使整个生命科学进入了分子生物学研究的新阶段;70 年代以后,发酵工程与遗传工程、细胞工程及酶工程紧密结合,使微生物工程处于生物工程技术的主角地位。在近代科学中,对人类贡献最大的一门科学当数微生物学了。微生物在医疗保健、抗生素药剂生产、工业发酵、农业科技、资源勘探、微生物采油等领域已经得到广泛应用。

（二）微生物与蛋白质组学

当前，适应基因组学和蛋白质组学的发展，微生物蛋白质组学应运而生。微生物蛋白质组学研究集中在两个方面，即一方面是各类模式微生物数据库的建立及一系列基础工作的开展，另一方面是病原微生物建库、病原物与疫苗的筛选与研制，可以说基础研究和应用研究的工作齐头并进、相互联系、互相推动。

据统计[3]：截至 2000 年 5 月，已经完成全序列测定的基因组有：古细菌 6 种，细菌 21 种，噬菌体 66 种，病毒 542 种（包括类病毒），细胞器 152 种，真核生物 3 种。而 3 年后，截至 2003 年 5 月 2 日，基因组全序列已完成的已经增至：古细菌 16 种，细菌 106 种，噬菌体 112 种，质粒 280 种，类病毒 40 种，病毒 875 种，细胞器 308 种，真核生物（包括完整染色体）15 种。蛋白质组全序列已完成或正在进行的有：真核生物 8 种［分别是植物拟南芥（*Arabidopsis thaliana*），秀丽线虫（*Caenorhabditis elegans*），果蝇（*Drosophila melanogaster*），*Guillardia theta*，人类（*Homo sapiens*），鼠（*Mus musculus*），酿酒酵母（*Saccharomyces cerevisiae*），裂殖酵母（*Schizosaccharomyces pombe*）］，古细菌 16 种，细菌 83 种。从这一组数据不难看出，微生物蛋白质组学的研究已经有了迅速发展的雄厚基础。

我们按模式微生物和病原微生物两部分来讨论蛋白质组学在微生物中的应用。

二、蛋白质组学在模式微生物研究中的应用

“In the Age of the Ome”，模式生物仍然是理解复杂生物学通路和人类疾病不可或缺的科研工具[4]。模式生物以人类基因组计划的提出为界分为两大类：前基因组时代的 BG 模式生物和后基因组时代的 AG 模式生物。AG 模式生物的选择主要取决于它们在物种进化、人类疾病研究、生物药物发展和工业应用等方面的潜能。典型的模式生物有以下的特征：适用正向和反向遗传学的各种方法，短的传代时间、体型小、生活条件简单且消耗合理，已有雄厚的研究基础（例如，PubMed 数据库中有关酿酒酵母的文献已达 50 000 篇），可以实现对生物学问题的简单化研究。原核的大肠杆菌和真核的酵母菌就是满足上述特征的“超级模型”（super model）。

（一）大肠杆菌蛋白质组学

大肠杆菌基因-蛋白质数据库 ECO2DBASE（*Escherichia coli* gene-protein database）是 Vanbogelen 和 Neidhardt 自 20 世纪 80 年代中期起就致力于构建的大肠杆菌蛋白质组数据库，截至 2001 年，已发展到第 6 版，包括了大肠杆菌的 10 种不同的外界环境。目前，大肠杆菌蛋白质组学的研究几乎涉及了蛋白质组学现有领域的方方面面。

在蛋白质表达及调控领域，加利福尼亚的 Doyle 等人发展了一条基于 96 孔板的纯化大肠杆菌蛋白质产物的高通量、高效、易操作的蛋白质组学生产流水线[5]。生物体对抗多种环境压力时除了调节特殊基因的表达之外，还可以整体性地同时调整一大批相关基因的系统表达，在这些系统中，了解最早的是热激反应中的热激蛋白家族，Rosen 和 Ron 总结前人在热激蛋白家族方面的工作，并首先以大肠杆菌等为模式生物将 2DEMS 的蛋

白质组学技术用于热激蛋白家族及其调控元件的研究[6]。

在结构蛋白质组学领域，Huang 以大肠杆菌的核糖体结合因子 A（RbfA）——一种冷休克适应蛋白为结构蛋白质组学的研究对象[7]。细菌或古细菌基因组编码的非膜结合蛋白质，经由大肠杆菌表达后只有一半是可溶的，这给功能和结构的研究带来极大的困难，针对这一问题，Savchenko 等人发展了结构蛋白质组学的一个战略思想[8]：核磁共振和 X 射线结晶技术同时研究不同物种的同源蛋白质，他们选取大肠杆菌和嗜热菌 *Thermotoga maritima* 的 68 对同源蛋白质做了尝试，取得了很好的结果。Ge[15]等人在研究大肠杆菌的硫胺生物合成酶系统成员 ThiS 和 ThiG，辅酶 A 生物合成酶系统的 CoaBC（phosphopanto-thenoylcysteine synthetase/decarboxylase ）等时，应用“top down”-MS 路线解释了它们与 DNA 预测序列的相对分子质量偏差在于末端剪切或翻译后的修饰，而在“bottom up”-MS 路线中无法解释这些较大的蛋白质与预测的偏差，提示电子俘获分离（electron capture dissociation，ECD）-top down FTMS 可以告诉我们有关较大蛋白质的细节结构等信息。

在膜蛋白组研究方面，Fountoulakis 和 Gasser 用弱去污剂（如 CHAPS）或强去污剂（如胆酸钠、去氧胆酸钠）的缓冲液抽提膜蛋白，2DE-MALDI-MS 蛋白质组学技术研究野生型大肠杆菌菌株和表达人细胞色素 P450 转化株的膜蛋白质组，共鉴别 394 种膜蛋白，其中 28%（112 种）符合预测软件的结果（其中预测含一个跨膜结构区域的有 100 种，含两个跨膜结构区域有 10 种，含三个跨膜结构区域有 2 种），使用强去污剂缓冲液抽提多鉴别出 20 种膜蛋白，这是目前最大的膜蛋白数据库之一。

在定量及非凝胶蛋白质组学方面，Yan 等用 2D-DIGE 结合 MALDI-TOF-MS 或 Q-TOF-MS 技术，鉴别了苯甲酸处理后的与对照组的大肠杆菌之间 179 种差别表达的蛋白质[12]。Gevaert 等人开发的联合分级对角线色谱（COFRADIC™）[9]是一种全新的非凝胶蛋白质组学技术，不同于现在流行的多维蛋白质鉴定技术（multidimensional protein identification technology，MudPIT）或是同位素代码亲和标签技术（isotope-coded affinity tag method，ICAT），它基于对角线色谱法（diagonal chromatography）[10]：抽提大肠杆菌 K 12 的全细胞蛋白质，胰蛋白酶酶解后，所有的肽段经反相柱（narrow-bore reverse-phase ZORBAX®300SB-C18 column）分成 48 个组分，分别用双氧水进行甲硫氨酸氧化反应，而后再用反相柱收集含蛋氨酸亚砜（methionine-sulfoxide，Met-SO）的肽段，用 LC-MS/MS 进行质谱分析，所得数据用 Mascot 和 Procorr™两种算法共识别 872 种（相当于大肠杆菌 E. coli K12 蛋白质组的 40%）含有甲硫氨酸的蛋白质，包括了可溶蛋白质和膜结合蛋白质、高丰度及低拷贝蛋白质。这比基于 2DE 的蛋白质鉴定技术灵敏 100 倍，并且可以完全实现自动化。更为重要的是，只要肽段中含有可修饰的氨基酸（如磷肽、氨基末端肽等），灵活运用肽段分类化学法，并结合胰蛋白酶催化反应时的水中氧同位素并入[11]就可以定量分析每一类肽段，鉴定出对应的蛋白质。Wang 和 Regnier 以大肠杆菌为对象，为对比质粒表达与否的差别蛋白情况，设计了如下的实验路线[17]：基于硫醇-二硫化物 thiol-disulfide 交换反应的共价色谱法选择胰蛋白酶酶解产物复合物中含半胱氨酸的多肽，再经质谱鉴定并按同位素标记定量：2，2′-联吡啶二硫化物打开二硫键，胰蛋白酶酶解，不同同位素标记（如 H、D）的琥珀酐酰基化，耦联了巯基基团的丙基硫氧嘧啶 Sepharose 胶通过二硫化物交换选择富集含半胱氨酸的多肽片段，洗脱下来后经反向液相色谱（reversed-phase

liquid chromatography, RPLC)分级，各个组分用 MALDI 等分别进行质谱鉴定并按同位素标记定量。

关于生物信息学对蛋白质相互作用图谱及生物学通路的预测方面：interacting domain profile pair (IDPP)方法[13]是通过序列及连接相似度将蛋白质结构域归类，这比仅仅基于蛋白质同源性[14]的方法更为有效。Wojcik、Boneca 和 Legrain 应用 IDDP 方法预测大肠杆菌、空肠弯曲杆菌的蛋白质相互作用图谱及比较大肠杆菌、空肠弯曲杆菌趋化性通路，并将预测的蛋白质相互作用图谱与酵母双杂交实验所得的相互作用图谱相互比较。以进化系统树的相似度为蛋白质相互作用预测的指示物[16]，成功预测了涉及大肠杆菌 2 742种蛋白质的 67 000 对相互关系。

另外，Wang 和 Regnier 将固定化胰蛋白酶柱与多维分离技术的联用[18]，20min 就实现大肠杆菌全蛋白质抽提物的酶解，在自动化蛋白质组学方面进行了有益的探索。

（二）酵母蛋白质组学

酿酒酵母（*Saccharomyces cerevisiae*）和裂殖酵母（*Schizosaccharomyces pombe*）是两种广泛使用的单细胞真菌类模式生物[4]，尽管它们有不同的生活方式，属于不同的进化分支，但它们都可以使用各种强大的分子遗传学手段，有极短的传代时间，这使得它们成为遗传学和基因组学的经典模式。AG 模式生物时代它们仍然是不可多得的真核模式微生物。酿酒酵母和裂殖酵母各自拥有 6 400 和 4 900 个基因，其中 30%是未知功能的，这些基因的功能是什么？它们如何表达调控？表达产物（尤其是蛋白质类产物）的亚细胞定位情况、相互作用、酶学等功能是怎样？这样庞大的任务无疑是模式微生物蛋白质组学必需打响的前哨战。蛋白质组学以系统论为基础，较为推崇高通量技术，如今在模式酵母中先行试用的有 5 种：①研究基因表达调控和蛋白质-基因相互作用[19,20]的 DNA 阵列技术；②研究蛋白质相互作用的酵母双杂交系统[21]；③研究蛋白质相互作用的蛋白质复合物纯化与质谱鉴定技术[22]；④通过 lacZ、绿色荧光蛋白质或同位素标签为报告者的亚细胞定位技术[23]；⑤实现了酿酒酵母 5 800 个可读框的克隆表达的蛋白质组芯片技术[24]。

基于这 5 种高通量技术的一系列蛋白质组学研究手段，酵母微生物组学取得了惊人的成绩。

Perego 等人对裂殖和酿酒两种模式酵母在真核细胞 DNA 修复、细胞周期调控、主要的信号通路、抗细胞毒素剂或抗肿瘤药物反应机制的模型等方面的贡献做了总结，突出介绍了它们的突变株用于药物筛选的工作[25]。Feldhaus 等人应用蛋白质组学领域的一种高通量的抗体分离技术来研究单链可变片段的抗体文库（single-chain variable fragments of antibodies scFvs，人的已有 10^9种 ），将 scFvs 克隆表达在酵母表面，磁珠法富集结合流式细胞术筛选（已经有噬菌体及核糖体展示文库），是迄今为止最大的人 scFvs 的酿酒酵母展示文库[26]。

在蛋白质复合物及蛋白质相互作用方面（相互作用组，interactome）的酵母蛋白质组学研究更是成果喜人。以酿酒酵母为模式，检验新的群集算法——分子复合物检测（molecular complex detection, MCODE），Bader 和 Hogue 收集整理 15 143 种酿酒酵母蛋白质间相互作用，涉及 4 825 种酵母蛋白质，相当于酿酒酵母蛋白质组的 75%，共预测了 209 个蛋白质复合物[27,28]。以生化法纯化结合质谱鉴定发现了酿酒酵母后期起始复合物（anaphase-promoting

complex,APC)的 11 个亚单位[29],Yoon 等人[30]采用串联亲和纯化 tandem affinity purification,TAP 和 direct analysis of the purified complexes by mass spectrometry (DALPC)技术在裂殖和酿酒酵母 APC 复合物都检测到 13 个亚单位,并且提示原来的技术可能只偏重了小亚单位的检出。Ohi 和 Gould 用裂殖酵母或酿酒酵母剪接体的必需元件 Cdc5p 或 Cef1p 分别共纯化出 30 种相互作用的剪接体复合物候选蛋白质[31]。另一个实验小组用 TAP 和 DALPC 技术研究裂殖酵母或酿酒酵母 Cdc5p/Cef1p 剪接体复合物,发现至少 26 种蛋白质组分[31]。酿酒酵母(*Saccharomyces cerevisiae*)是最重要的真核模式生物之一,对其蛋白质组的研究,对于阐明人类基因组的功能有着重要的方法学与同源提示作用。2002 年,几个研究组在《自然》和《科学》上相继发表了对酿酒酵母的大规模功能蛋白质组研究的初步结果:高通量质谱检测蛋白质复合物方法(high-throughput mass spectrometric protein complex identification,HMS-PCI)是基于 FLAG-标记的酵母可读框的重组体克隆(recombination-based cloning),免疫共沉淀后用 anti-FLAG 的琼脂糖树脂捕获蛋白质复合物,而后 1D-SDS-PAGE 分离,LC-MS/MS 鉴定,生物信息学分析数据,结果用 493 个诱饵(约占酵母蛋白质组的 10%)共检测到 3 617 相互关联的蛋白质,涵盖了酵母蛋白质组的 25%[32]。另一个研究小组用 TAP 方法标记了约 1 700 个蛋白质,其中1 143个与人类相关,经质谱分析鉴定了 232 个蛋白质复合物和 344 个蛋白质可能的细胞功能[33]。Tong、Drees、Nardelli 和 Bader 等人,应用基于噬菌体展示配基共有序列的算法预测与大规模酵母双杂交相联合的战略,仅利用酵母的 SH3 结构域就分别鉴别了涉及 206 种蛋白质的 394 对相互作用和有关 145 种蛋白质的 233 对相互关系[34]。

定量蛋白质组学中酵母也发挥着作用。例如,ICAT 标记葡萄糖或半乳糖为碳源的酿酒酵母蛋白质,在不同质量标签标记的同种蛋白质在 2DE 中共移行,MALDI-TOF-MS 检测了 13 种这样的共移行蛋白质(peroxisomal alkyl hydroperoxide reductase,enolase 2,aldolase,phosphoglycerate kinase,glyceraldehyde-3-phosphate dehydrogenase 3 等)[35]。Smith 等人[36]用傅里叶变换离子回旋加速共振质谱 FTICR 或串联质谱分析总蛋白酶解产物,以酿酒酵母和球菌 *Deinococcus radiodurans* 为研究对象,也进行了蛋白质定量方面的尝试。

另外,磷酸化蛋白质组学方面,Chen 等人使用多种基于质谱的方法分析酿酒酵母细胞分裂后期的重要调节者 Net1 的磷酸化位点谱图,包括金属螯合柱、碱性磷酸酶处理、基质辅助激光解吸源后衰变、多维电喷雾质谱,这些方法相互补充,但任何一种都无法找到所有的磷酸化位点。使用的蛋白质样品是酿酒酵母重组 Net1 菌株和突变体 Net1 N19m 菌株(19 个丝氨酸或苏氨酸位点突变为丙氨酸)[37]。结构蛋白质组学方面,耶鲁大学组建的一个结构蛋白质组学的数据库 SPINE,在其建立伊始就收录了 123 种酿酒酵母蛋白质的结构信息[38]。

因篇幅所限,我们仅以模式微生物中最具代表性的大肠杆菌和酵母来讲述现今极具发展规模的模式微生物蛋白质组学。

三、蛋白质组学在病原微生物研究中的应用

在已知的 10 万多种微生物中有少部分为病原微生物[39]。病原微生物可以参照微生物的分类方法分三大类:①非细胞型的病毒是最常见的致病微生物;②原核细胞型的细

菌、枝原体、衣原体、立克次体、螺旋体、放线菌的某些种类；③真核细胞型的真菌、原生动物、藻类中也含少量致病微生物。这些病原微生物引起的感染性疾病一直是人类健康的大敌，尽管在抗生素和抗病毒药物问世以后，这类疾病对人类的威胁已经减小。但是，由于新的微生物的出现（如最近祸及全球的非典型性肺炎的致病病毒是一种崭新的冠状病毒），以及一些已有的病原微生物产生了抗药性，感染性疾病仍然是全球范围内主要的疾病之一。

自 1995 年，第一个致病菌基因组测序完成以来，基因组学在抗感染药物研究领域中做了大量的基础工作，但遗憾的是，截至 2002 年基因组学并未发现一种全新的抗菌试剂[50]。因而，蛋白质组学、生物信息学等的发展是必需的。对这些致病微生物的蛋白质组学研究，对于了解其毒性因子、抗原及疫苗的制备非常重要。此外，对疾病的诊断、治疗和预防也同样重要。现已获得的病原微生物基因组学等工作的丰硕成果为其蛋白质组学研究提供了有利条件。目前，关于病原微生物蛋白质组学研究主要集中在以下几个层面[73]：①作为病原微生物蛋白质组学研究基础的模式微生物在不同生长条件下的表达谱的研究；②同一种病原微生物的不同菌株之间的蛋白质组的比较；③同一种病原微生物的不同生理状态的蛋白质组学研究；④对病原微生物进行“亚蛋白质组”（Sub-proteomics）分析；⑤研究宿主和微生物的相互关系，寻找新型疫苗的靶分子。第①方面的工作可以参阅本节的概述及蛋白质组学在模式微生物研究中的应用两部分，在此简单介绍②～⑤方面的工作。

（一）同一种病原微生物的不同菌株之间的比较蛋白质组学

部分病原微生物的不同菌株具有不同的致病性，研究这些菌株之间的蛋白质组学的差别，显然有助于了解该微生物的致病机制以及相关药物、疫苗的开发。

螺旋体（*Microbes spirillum*）是一类既有细菌某些特点又有原生动物某些特点的微生物，它们当中很难找出对人类有益的种类，那些生活在人和动物体内的螺旋体会使它们的寄主患上许多种疾病。其中的钩端螺旋体属（*Leptospira*）中的许多种类更为凶恶[39]。本实验室针对钩端螺旋体的两个不同致病性的菌株以及它们在不同生长条件下的蛋白质组进行了比较，工作中使用了双向凝胶电泳、质谱鉴定技术、液质联用技术和同位素标记定量技术，对这些菌株进行了大量细致的分析。结果表明，这些菌株在蛋白质组水平上有许多重要的差别，为进一步的药物和疫苗开发提供了条件。这一成果即将发表，可参见相关的基因组学工作[40]。

幽门螺杆菌感染与慢性胃炎、胃和十二指肠溃疡有关，利用 2DE-MALDI-TOF 蛋白质组学技术研究 4 个幽门螺杆菌临床隔离群，对比差别蛋白表达图谱，4 个临床隔离群明显按地区来源分成了两组，这表明蛋白质组学研究可以作为致病微生物临床隔离群区分的可靠参数之一[46]。

（二）同一种病原微生物的不同生理状态的蛋白质组学研究

疟疾是现代三大瘟疫（结核病、艾滋病、疟疾）之一，2002 年 10 月英国《自然》杂志公布了疟疾的致病微生物——恶性疟原虫的基因组，同一期《自然》杂志也刊登了多国学者在恶性疟原虫的蛋白质组的研究中取得的成果：针对疟原虫的各个生理状态下的蛋白质

组表达情况进行分析，全面了解了疟原虫的生命周期中的蛋白质组[44,45]，即应用多维蛋白质识别技术 MudPIT 研究恶性疟原虫 clone 3D7 在孢子期、裂殖期、滋养体期和配子体期的蛋白质组，共鉴别 2 400 种蛋白质；或用 NanoLC-MS/MS 分别鉴别了恶性疟原虫无性生殖期 714 种、配子体 931 种和配子期 645 种共 1 289 种蛋白质。2003 年 PlasmoDB (http://PlasmoDB.org) 提供了恶性疟原虫的全基因组的信息及注解。

近来，Reid 等人对链球菌的不同生长状态时的蛋白质组变化进行了研究，并和基因表达进行了比较，对 4 种新的抗原分子进行了分析，这些分子是潜在的药物和疫苗的靶分子[43]。

Demirev 等人自编程序推测革兰氏阴性菌——幽门螺杆菌 26 695 株潜在的蛋白质标志物的翻译后修饰，可靠性提高了至少一个数量级[42]，这是生物信息学在病原微生物蛋白质组学中应用的又一有益尝试。Lock 等人使用 2DE-MS 鉴别了幽门螺杆菌 NCTC 11 637的 93 个高丰度蛋白质，对应 35 个基因，每基因 2.7 个蛋白质产物[51]。

Kahng 等人开展对不动杆菌的碳源分解代谢研究[49]：生长在琥珀酸盐或 phydroxybenzoate 培养基不动杆菌属的 *Acinetobacter lwoffii* K24（可以分解磺胺药物前体 aniline），对比它们经 Q-Sepharose 柱分离后的 2DE 蛋白质图谱，用 N 端测序（ Edman Degradation）和内部测序（ESI-Q-TOF mass spectrometry）鉴别了差别表达的两种原儿茶酸-3,4-二加氧酶亚基 pcaH 和 pcaG，它们都与 p-hydroxybenzoate 的分解代谢有关，可能是该菌株耐药性产生的源泉所在。

一类胞内革兰氏阴性菌——衣原体属可引起人的急性或慢性的疾病，肺炎衣原体和沙眼衣原体的基因组计划使我们对衣原体致病机制及潜在毒性有了更深的认识。Shaw 等人[62]对此进行了较为深入的工作并补充了沙眼衣原体（*Chlamydia trachomatis*）蛋白质组：2D-SDS-PAGE 结合串联质谱对比胞外传染性原生体（extracellular infectious elementary body, EB）和胞内网状体（intracellular reticulate body, RB），发现一个全新的 204bp 的开放阅读框，编码的蛋白质 7kDa reticulate body protein（Rbp-7）无其他种属的同源蛋白，肺炎衣原体（*Chlamydia pneumoniae*）中也未找到。深入的工作显现了 Rbp-7 的动态表达，它在 EB-RB 循环中产生，只存在于胞内网状体内并迅速降解，提示其潜在的致病机制。另一个研究小组[63]用^{35}S标记的沙眼衣原体蛋白质，结合 2DE-MALDI-TOF 对比体外 γ 干扰素处理或未处理的持续感染了肺炎衣原体 Hep-2 细胞，发现体外 γ 干扰素处理后肺炎衣原体的一些病原体蛋白质（如 Major Outer Membrane Protein (MOMP), Heat Shock Protein 60 (Hsp-60/GroEL)、Proteins With Functions In DNA Replication (GyrA), Transcription (RpoA, PnP), Translation (Rrf), Glycolysis (PgK, GlgP), Type III Secretion (SctN)）表达无明显的上升，提示 γ 干扰素的药物作用机制，进一步的研究将围绕该衣原体生长周期中 γ 干扰素的作用展开。

支原体（*Mycoplasma*）是一类缺乏细胞壁的原核细胞型微生物，它既不同于细菌，也不同于病毒，是介于细菌和病毒之间，能在无活细胞的人工培养基上生长繁殖的最小微生物。支原体的种类繁多，造成的危害非常广泛。目前，研究较多的是肺炎支原体（*Mycoplasma pneumoniae*, MP）和生殖器支原体（*Mycoplasma genitalium*, MG）。Regula 等人[64]应用不同的 pH 梯度、不同胶浓度、不同的缓冲液系统进行 1D 或 2D-SDS-PAGE, LCQ 或 MALDI 鉴别了 350 种蛋白质，对比原于 DNA 序列的基因产物预测，350 种蛋白

质对应 224 个基因，初步建立了肺炎枝原体蛋白质组数据库；Ueberle 等人[65]则在此基础上，将差速离心、阴离子或阳离子交换层析、肝素配体的亲和层析、固定的抗生蛋白链菌素分离生物素结合蛋白质等新方法用于肺炎枝原体蛋白质组数据库的扩充，实现了从 224 个基因产物到 305 个基因产物的分离鉴定。生殖器枝原体是最小的柔膜细菌、最小的自养型单细胞生物，基因组大小仅仅为 580kb，只有 480 种基因表达产物，是研究自由生活细胞所需最低新陈代谢及蛋白质组学研究的极好模型。4 个相互重叠的 pH 梯度进行 SDS/PAGE，发现 427 种蛋白质，质谱鉴定了其中 158 种。指数生长期过后，生殖器枝原体蛋白表达谱的变化：DnaK，GroEL，DNA 回旋酶和细胞黏连附属蛋白高表达，一些核糖体蛋白质低表达[66]；作为补充工作，Balasubramanian[67]将生殖器支原体的所有 483 个可读框分类，研究其中 25 个未知同源结构及功能可读框，用圆二色性（circular dichroism，CD）等方法分析它们基因产物的生物物理学及三维结构的特性，并建立了生殖器支原体的蛋白质组数据库：http://bioinfo.mbb.yale.edu/genome。

20 世纪 90 年代，随着氟康唑和伊曲康唑两价目三唑类通过抑制甾醇生物合成的抗真菌药物的出现，给机会致病真菌（如二态的白色念珠菌）感染的治疗带来了希望，但随后念珠菌属产生了许多的耐药菌株。新近研究发现，Echinocandin 类化合物是一类杀真菌剂，通过抑制真菌细胞壁的 α-葡萄糖的合成而起作用，而且对三唑类真菌药物的耐药菌株起作用。为了进一步研究这两类抗真菌药物，Bruneau 等人[68]对比 mulundocandin、氟康唑和伊曲康唑三种杀真菌药处理所产生的白色念珠菌 2-DE 差异蛋白质图谱后认为，这两类抗真菌药物有蛋白质组水平共同的作用机制。最近，Lilly 实验室的 Echinocandin 及 Merck 公司的 Pneumocandin 已在进行 I 期和 II 期的临床试验。

隶属于孢子纲 Sporozose 弓形虫科 Toxoplasmatidae 的刚地弓形虫或称鼠弓形虫（*Toxoplasma gondii*），其主要寄生部位为有核细胞。Cohen 等人[70]以对数生长期的鼠弓形虫为研究对象，用 pH 4～7 和 pH6～11 可分离到 1 000 余种多肽，用多重的更窄的 pH 梯度胶可分离到 3 000～4 000 种多肽，再用鼠弓形虫等原生动物的 EST 数据库解释 MALDI 质谱所得肽质量指纹图谱的可行性与可靠性，这是蛋白质组技术与生物信息学技术的又一次合作。

（三）病原微生物的亚蛋白质组学

病原微生物的某些蛋白质对药物和疫苗研究有着重要意义，包括分泌蛋白质、膜蛋白质和与人体免疫相关的蛋白质，等等，因此对这些“亚蛋白质组”的研究也成为热点。例如，膜蛋白质对于病原微生物研究有着重要意义，往往是致病机制中重要的抗原分子。

Antelmann 等人对枯草芽孢杆菌的细胞壁进行了蛋白质组水平上的研究，分析了细胞壁上蛋白质的稳定性[52]。此外，在疫苗研究中，也有对致病微生物的膜蛋白质进行“血清蛋白质组”分析，这部分内容将在下一小节详细[53,54]介绍。另外，有关芽孢杆菌 *Bacillus sp* K-1 菌株分泌蛋白的蛋白质组学的初步研究[47]：用 2-DE 和 N 端测序法鉴别了 8 种分泌蛋白。

革兰氏阴性菌空肠弯曲杆菌（*Campylobacter jejuni*）感染称为弯曲杆菌病，弯曲杆菌在胆管中生长，小肠上端微需氧环境适宜本菌生存和繁殖，造成空肠、回肠和大肠组织损伤，通过肠黏膜侵入血液。空肠弯曲菌感染的患者，临床表现发热、腹泻、呕吐和肌肉痛。

重症病例也有痢疾样血、黏液便,常误诊为急性溃疡性肠炎。Young 等人 ESI-MS/MS 质谱分析空肠弯曲杆菌蛋白质 PEB3 的糖基化修饰:大豆植物凝集素亲和层析、2D-PAGE、链霉蛋白酶酶解、凝胶过滤、MALDI-TOF-MS 和 nano-NMR 等技术联用,鉴定了至少 22 种 N-linked glycan 糖蛋白并分析了聚糖结构[48]。

Graves 等人[69]用蛋白质组学方法结合亲和层析技术(多种喹啉类药物或含嘌呤环结构的类似物挂在 Sepharose 柱上观察蛋白质复合物中的相互作用的蛋白质)探究喹啉类药物[如(抗疟药)羟氯喹、伯氨喹的]的作用靶点及可能机制。鼠、人红细胞蛋白质组中发现 ALDH1 醛脱氢酶 1 和 QR2 苯醌还原酶 2 两种喹啉结合蛋白质,但在恶性疟原虫蛋白质组中未发现这两种蛋白质。

Rout 和 Field[71]提纯锥虫属 *Trypanosoma brucei* 的细胞核(包括亚核组分),应用电泳、电镜、免疫印迹、免疫电子显微镜等技术,发现并证明一种新的核纤层蛋白 NUP-1,在 bloodstream 和 procyclic stages 差别表达。锥虫外壳蛋白质包括了多种亚型的脂质体-锚蛋白和糖蛋白等。最近免疫化学、生物化学、质谱测量法等多学科技术综合应用,得到了锥虫外壳蛋白质的更多信息,也可以用来解释带菌者采采蝇与寄生虫非洲锥虫之间动态的相互作用[72]。

(四)研究宿主和微生物的相互关系,寻找新型疫苗的靶分子

长期以来,微生物引起的疾病一直严重威胁着人类健康,研究宿主和各种病原微生物的相互关系,宿主对病原体的初级应答反应、体液免疫反应、细胞免疫反应,是发现新的病原物、抗原类物质、构建新疫苗和筛选新药物的重要手段。

疫苗是人类用于治疗微生物感染类疾病的有效方法,但是疫苗的获得是非常困难的,常常需要使用灭活的病原体或者其他危险的抗原。以上这些物质对于人体是很危险的,现在蛋白质组学技术等后基因组技术的出现为更加理性的设计疫苗提供了可能。以双向凝胶电泳、生物质谱、生物信息学、加上免疫学的技术(即所谓“血清蛋白质组分析”——SERPA)[56]为代表的几大技术对于推动蛋白质组学在疫苗学研究中的应用起到了关键作用。其中,SERPA 技术已经应用到了多项疫苗研究中,其中不乏成功的例子[54,57~59]。疫苗学的发展显然是病原微生物蛋白质组学的热点之一,有多个领域的工作在此交汇,cDNA 文库技术、重组 DNA 技术和基因、蛋白质芯片技术加上蛋白质组技术一起开展工作。在这里已经形成“免疫组”(immunome)[60],并将进一步发展,也许不久免疫组的各项技术将成为免疫学的常规技术,为理性地设计疫苗分子提供思路。甚至通过对多种致病微生物的免疫组研究,可能会发现它们的共同点,就是说可能就此发现对多种微生物感染有效的疫苗。同样,这些概念在癌症的疫苗研究中也正发挥着重要的作用,对各种癌症的免疫组研究正在进行。这些工作中,关于致病微生物的研究是令人关注的,下面以金黄色葡萄球菌、幽门螺杆菌和包柔疏螺旋体为例介绍作一简介。

金黄色葡萄球菌(*Staphylococcus aureus*)是葡萄球菌中的一种,也是最危险的一种。大约 1/3 的人会在不知不觉中感染金黄色葡萄球菌,但是在医疗过程中,有可能在外科手术和脓肿等情况时,感染这种微生物,形成菌血症、脓血症、心内膜炎和肺炎等,造成严重后果,危及人的生命。更严重的是,一些耐药性金黄色葡萄球菌已经出现,其中的一种甚至连万古霉素都无效[61]。Vytvytska 等人对耐药性金黄色葡萄球菌进行了血清学蛋白质

组分析[54],发现了几种高度血清敏感性的抗原,它们可能作为疫苗的候选抗原分子。更早的血清学蛋白质组学分析(SERPA)思路来源于 Klade 等人的肿瘤免疫学工作,这个思路运用到微生物研究中时,结合了血清学、双向凝胶电泳技术、免疫印迹法和胶内酶解蛋白质质谱鉴定技术。Vytvytska 等人的工作分为三个部分,基本上反映了血清学蛋白质组学分析(SERPA)的思路。第一部分是高滴度人血清的获得,这一部分工作是免疫学的内容,作者将所有的血清样品分成健康人、被感染者和作为对照组的婴儿,使用 ELISA 技术测定各样品对金黄色葡萄球菌抽提物和重组的细菌表面分子的 IgG 反应水平,从中挑选了免疫反应最强的样品用于下一步的工作。第二部分是优化细菌生长条件,表达尽可能多的抗原分子。第三步是细菌膜蛋白质的制备,双向凝胶电泳,然后使用高免疫反应的血清样本做免疫印迹,找到与血清中抗体结合的蛋白质点,最后将平行操作的制备级双向凝胶电泳胶上相应的点切下做鉴定。用这种方法找到的抗原不仅仅可以作为疫苗的目标分子,还可能为新药的研制提供信息(比如,有些抗原分子只在耐药菌株中出现)。

另一项出色的工作是针对幽门螺杆菌的研究[53]。这项工作所用的技术与上述相似,研究总共发现并鉴定了 32 个与阳性血清特异性结合的蛋白质点,其中 9 个是新报道的蛋白质,其余也被其他研究成果所验证。在这 9 个蛋白质中有 3 个和阳性血清高度特异性结合,被认为是很有希望的疫苗候选分子,进一步的工作已经展开,其他的阳性蛋白质也可能成为临床诊断标记分子。

莱姆病(Lyme disease)病原为包柔疏螺旋体,由蜱传播,症状有发热、头痛、慢性移行性红斑,可表现为关节炎、脑膜炎、神经炎、脊髓炎和心肌炎等。这一疏螺旋体属至少包括了 5 种不同致病性的菌株。奥地利 Jungblut 等人选取 *B. garinii* BITS 菌株为研究对象,应用 2-DE-MALDI-MS 对比它在杂交了较早(红斑迁徙期)和较晚(出现关节炎、肢皮炎等症状)临床时期的病人血清的差别蛋白质组,鉴定了包括疾病时期特异性蛋白质在内的 20 种抗原类物质[41],这些抗原均是潜在的可用于特异性诊断的标志物。

Chakravarti 等人[55]也为研究者们提供了一个联合基因组学、转录组学、蛋白质组学和生物信息学,用三种不同的路线寻找疫苗候选者的方案(图 23.1),应用这一方案的研

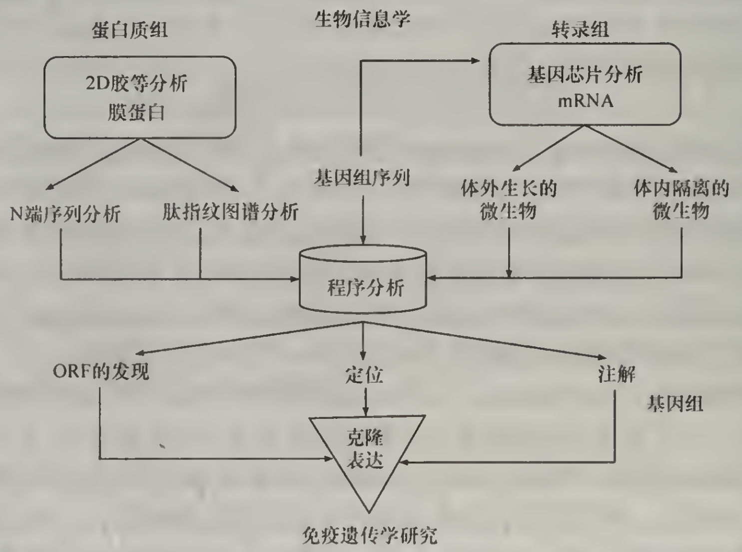


图 23.1　疫苗候选者的方案

究工作有:生物信息学预测流感嗜血菌 Rd 基因组中一个潜在的疫苗候选者——外膜脂蛋白 P6;分别用 1-DE-PAGE-N-端测序寻找与单抗相互作用的蛋白质和 2-DE-MS 的两条路线筛选幽门螺杆菌潜在的疫苗候选者。

四、展　　望

以上简单列举了微生物蛋白质组学在模式微生物和病原微生物两大方面的研究成果,这两方面是蛋白质组学在微生物中应用的重点。

如今,各种综合的或是专门的收录微生物蛋白质组的数据库日益完善。例如:为欧洲细菌蛋白质组计划(European bacteria proteome project,EBP)而建立的 http://www.mpiibberlin.mpg.de/2D-PAGE 数据库;欧洲致病(菌)微生物数据库(http://www.mpiibberlin.mpg.de/2D-PAGE)2002 年 1 月建库,收录的第一批数据包括结核分枝杆菌(*Mycobacterium tuberculosis*)、幽门螺杆菌(*Helicobacter pylori*)、包柔疏螺旋体(*Borrelia garinii*),土拉热弗朗西斯菌(*Francisella tularensis*)、肺炎衣原体(*Chlamydia pneumoniae*)和肺炎枝原体(*Mycoplasma pneumoniae*)等重要病原微生物的 2D-PAGE 图谱;瑞士生物信息学学院的 ExPASy (Expert protein analysis system, http://us.expasy.org);恶性疟原虫的 PlasmoDB(http://PlasmoDB.org);大肠杆菌基因-蛋白质数据库 ECO2DBASE。

近年来,WHO 越来越重视感染性疾病对人类健康的影响,各国政府也适时加大了资金投入并加强了病原微生物蛋白质组学的跨国跨地区的协作。国际协作的一个成功范例是,由美国国立卫生研究院的主席 Harold Varmus 于 1997 年倡议并实施的恶性疟原虫蛋白质组学的研究,他们还将以每年 1 亿美元的投资继续后期更为艰难且更具挑战性的功能蛋白质组方面的研究工作。

当然,模式微生物和病原微生物这两大方面并不能涵盖这一领域的所有研究内容,同时这两方面还有很大的发展空间,比如,原核细胞型的放线菌 *actinomycetes*、立克次氏体属(*Rickettsia*)的蛋白质组学研究工作尚未起步,植物致病菌、食品微生物、极端微生物的蛋白质组学研究刚刚开展。相信随着更新的蛋白质组学技术设备的发展,更密切的国际协作的进行,微生物蛋白质组学的研究将逐步走向深入。

第二节　蛋白质组学在肿瘤研究中的应用

随着人类基因组计划的完成,分子医学的研究内容正从基因组转向蛋白质组。在癌症研究中,一个关键性的问题是如何将蛋白质组技术用于基础和临床研究。在后基因组时代,蛋白质组技术可以用来发现肿瘤早期检测和诊断的分子标记,药物作用的靶分子和评估临床医疗的效果。本节将对蛋白质组技术及其在不同肿瘤研究中的应用加以总结概括。

肿瘤是威胁人类健康的重大疾病,它的发生和发展是一个多步骤、多因子控制的复杂过程。在环境和遗传等多方面的因素作用下,细胞内的原癌基因被激活和抑癌基因功能的丧失,细胞的正常生长和凋亡失去控制,而发生肿瘤;同时,随着肿瘤的恶性程度的增加,基因突变的进一步积累,癌细胞开始向周围组织入侵并穿透基质进入血液或淋巴循环

系统，在远处其他的组织器官中定居和生长，发生转移。对于临床的肿瘤学家来说，最迫切想知道的两个问题有两个。①肿瘤的诊断。肿瘤的早期诊断是一个非常重要的问题，肿瘤诊断得越早意味着临床的医疗措施会更有效。复旦大学中山医院肝癌研究所汤钊猷提出的小肝癌的概念就是一例[74]。除了早期诊断之外，目前临床上肿瘤的分级归类主要还是靠形态学观察，分子水平的鉴定目标不多，其手段也比较缺乏。②肿瘤的治疗。目前临床上对很多肿瘤最有效的治疗措施是手术切除，有效的肿瘤药物并不多，特异性的药物更少，原因是药物作用的靶分子很有限，某些肿瘤目前还没有很好的药物来治疗。因此，发现新的药物作用靶分子，从而开发出临床上行之有效的肿瘤药物也是亟待解决的问题之一。

生物技术的发展使人们对解决以上问题充满信心。随着 2001 年人类基因组全序列的完成和发表[75]，21 世纪分子医学的研究也已从基因组时代进入功能基因组时代，大量基因组序列信息的获得使人们的注意力正集中到探索诸如细胞生长、分化、凋亡和癌症发生等各种生命过程中基因组的功能。在基因组水平上分析不同条件下基因的表达，如基因表达的系列分析（serial analysis of gene expression，SAGE）[76]和 DNA 微阵、芯片（DNA microarray，chip）[77,78]等，是基因组功能研究的重要方面。然而，多数情况下 mRNA 水平与其被翻译的蛋白质之间并不存在较好的相关性[79]。基因组全序列的获得对蛋白质组的研究起了必不可少的推波助澜的作用，但蛋白质组又是一门相对独立的学科。它主要是在细胞整体水平上研究蛋白质的属性（表达水平、翻译后修饰、相互作用等），并由此在蛋白质水平上获得对疾病过程、细胞生理生化过程和调控网络的广泛而完整的认识，它是功能基因组研究的重要方面。“蛋白质是理解细胞功能和疾病过程的核心，没有蛋白质组学的配合和努力，就不能完全实现基因组学的成果。”[80]

目前蛋白质组学技术在肿瘤研究中已经得到越来越多的应用，并被认为是跨越基因组与临床应用之间鸿沟的桥梁。技术在蛋白质组学研究中扮演着至关重要的角色，蛋白质组学的发展直接受技术的推动和限制。从技术和研究目的角度，肿瘤蛋白质组学可分为表达蛋白质组学和功能蛋白质组学两大类。

在表达蛋白质组学研究中，通过特定肿瘤组织、体液或血清中的蛋白质的表达不同，可以识别肿瘤发生的分子标记，这一点在肿瘤的早期诊断和临床治疗方面非常重要。表达蛋白质组学所使用的技术手段主要是双向凝胶电泳（2-DE）、多维蛋白质鉴定技术（MudPIT）和质谱（MS），以及表面增强的激光解析离子化-飞行时间-质谱（SELDI-TOF-MS），特殊的新方法还包括将组织切片直接做基质辅助激光解析离子化-飞行时间-质谱（MALDI-TOF-MS），称为成像质谱（imaging MS）[81]。

功能蛋白质组学主要分析蛋白质-蛋白质或蛋白质- DNA/RNA 的相互作用和蛋白质的翻译后修饰。它对于研究细胞信号转导通路中蛋白质复合物的组成，以及这些复合物中组成蛋白质是如何相互作用并影响细胞内外信号会有所帮助。功能蛋白质组学的研究将会增加我们对癌细胞的生物形成过程的了解，并对开发新的肿瘤药物产生重要影响，引导我们发现基于蛋白质的肿瘤特异性靶分子。它所使用的技术手段主要是免疫沉淀法（IP）和 MS，以及蛋白质芯片和酵母双杂交。

蛋白质组学技术的比较以及它们在生物分子标记发现和肿瘤早期检测中的应用见表 23.1。

表 23.1　蛋白质组学技术的比较以及它们在生物分子标记发现和肿瘤早期检测中的应用

项目	酶联免疫吸附法(ELISA)	双向凝胶电泳	多维蛋白分析技术(MudPIT)	蛋白质组图谱诊断	蛋白质微阵列(protein array)
灵敏度	最高	总体较低,尤其是低丰度蛋白;受检测方法限制;LCM可以增加特异性	高	中等,对于高分子质量蛋白灵敏度会降低;新的质谱仪将提高灵敏度	中/高
生物分子标记的直接鉴定	否	是	是	否,新的 MS 仪器可以	和 MS 联用时可以
用途	检测在体液和组织中单个已知的分析物;临床分析的金标准	可用于生物分子标记的发现,但不是早期检测的直接方法	可以检测和鉴定潜在的生物分子标记	体液和组织中诊断图谱分析;潜在生物分子标记的鉴定	同时对多个分析物的分析
通量	中	低	很低	最高	高
优缺点	可机械化;临床分析中广泛使用;需要有广泛有效性的抗体来检测和确认效果;与直接发现不能互补(受检测方法的限制)	在临床应用之前必须对所有鉴定结果进行测试和确认,已被使用且证明是有效的方法。重复性好,和荧光染料结合可提高定量能力	与双向凝胶电泳相比有更高的灵敏度,对于生物分子标记的鉴定有更高的覆盖率	蛋白质的名称对于诊断图谱分析来说不是必需的;需要解决重复性和有效性的问题;若与生物信息学手段结合也许可以为临床化学领域带来变革	方法灵活:可以在一个样品里分析多个分析物或者在许多样品里分析单个分析物;需要事先了解被测分析物,受限于抗体的特异性和灵敏度;需要使用放大标签检测系统

蛋白质组学研究现已在多种人类肿瘤组织或细胞系(如肾癌、前列腺癌、卵巢癌、肺癌以及肝癌等)中开展,并取得了一定的成果。目前该领域主要集中于发现新的肿瘤分子标记以及对肿瘤发生学的研究。

为便于分析,以下将蛋白质组技术在各个不同肿瘤研究中的应用加以归类,见表 23.2。

表 23.2　蛋白质组技术在各个不同肿瘤研究中的应用

研究对象和目的	材料	方法与结果	参考文献
妇科肿瘤(gyneocologic malignacies)			
乳腺癌(breast cancer)			
蛋白质组方法区分两种正常的乳腺细胞	缩乳术后得到的正常的乳腺腔细胞(luminal)和肌上皮细胞(myoepithelial)	2-DE:找到大约 170 个蛋白质丰度变化在 2 倍以上;MALDI-TOFMS: 33 个蛋白质在乳腺腔细胞上调,18 个蛋白质在肌上皮细胞中上调;另外 134 个蛋白质无显著差异	[82]

续表

研究对象和目的	材料	方法与结果	参考文献
α烯醇化酶	人乳腺癌细胞系 MDA-MB231	2-DE、HPLC、MALDI-TOF MS 线性模式；用 PSD mode 和 LC/ESI-ITMS 测序，后者所得数据较好；找到 2 种新的共电泳的变异体：乙酰化修饰的 Asn-153Asp 和 Ile-152Asp/Asn-153Ile	[83]
与 MLK2 的 SH3 结构域相结合的蛋白质	人乳腺癌细胞系 MDA-MB231	从细胞裂解液中以 GST-MLK2N-GSH-Sepharose beads 作为亲和配体分离蛋白质，蛋白质洗脱后经电泳和胶内酶解，用 LC/ESI-MS/MS 鉴定出 HMG1	[84]
与乳腺肿瘤恶性生长相关的蛋白质	永生化的正常乳腺细胞株 MCF-10A 和两种相应的癌变株 MCF-10 Cala 和 Cald	HPLC、1-DE、电印迹和 MALDI-TOF MS 发现 p53 的突变体，c-src，c-myc 和 H-ras 在癌变细胞株中表达量升高	[85]
与乳腺肿瘤恶性生长相关的蛋白质	正常组织、恶性肿瘤组织和良性肿瘤组织	2-DE 和 MS，在恶性肿瘤中鉴定到 7 个蛋白质的表达量有 2～7 倍的升高；良性肿瘤中鉴定了 5 种蛋白质，其中 DJ-1A 只在乳腺纤维腺瘤组织中表达	[86]
鉴定 BclXL 以及和它相关的蛋白质	人胚胎肾细胞系 293T	将 BclXL 的质粒在细胞中表达，经免疫沉淀，用 LC/ESI-MS/MS 和 SEQUEST 鉴定到的蛋白质包括人蛋白质 CLARP、一种蛋白质激酶、一种磷酸酶等	[87]
伴侣分子 14-3-3ä	乳腺癌细胞系 MCF-7 和 MDA-MB-231，乳腺癌组织和正常乳腺上皮细胞	2-DE 和 MALDI-TOF MS 检测各种 14-3-3 蛋白质；定量研究显示仅 14-3-3ä 在癌细胞中表达量下调至原来的 1/10	[88]
α-微管蛋白同种型	乳腺癌 MDA-MB-231 细胞系	2-DE 和 MALDI-TOF MS(负离子模式)、ESI-ITMS 进行分析，根据 CID 及 PSD 碎片图搜索不同的数据库．找到一种以前未知的 α 同种型并进行测序	[89]
在血清中寻找与乳腺癌相关的蛋白质	正常人、乳腺癌患者和良性乳腺疾病患者	用 SELDI-TOFMS 发现一个 28.3×10^3Da 的蛋白质在所有晚期已转移的肿瘤中表达，而在其他各期肿瘤中少部分表达，在正常对照组和良性肿瘤中基本不表达	[89]
卵巢癌 (ovarian cancer)			
与肿瘤恶性生长相关的蛋白质	正常卵巢组织和卵巢癌组织	2-DE、MALDI-TOFMS 和 ESI-ITMS，发现 7 个蛋白质在癌组织中上调，其中 5 个只在癌组织中表达	[85]
肿瘤发生的分子标记	正常人血清和卵巢癌患者血清	SELDI-TOF MS 和生物信息软件，通过对 50 组正常对照和肿瘤患者的 SELDI-TOF MS 结果分析，训练过的软件能鉴定出全部 50 例卵巢癌患者，63/66 例的非卵巢癌患者，灵敏度达 100%，专一性达 95%，阳性预测值为 94%	[91]
泌尿科肿瘤 (genitourinary malignacies)			
肾癌(renal carcinoma)			
α干扰素和白介素 4 刺激下的肾癌细胞生长变化	人肾癌细胞系 ACHN	2-DE、Edman 降解法测序和 MALDI-TOF MS with PSD，检测到几个组成型蛋白质，在细胞因子 α 干扰素和白细胞介素 4 刺激下，原肌球蛋白，HSP27，Mn SOD，GST ð，蛋白激酶 C 抑制剂等 5 种蛋白质表达有不同程度上调	[92]
膀胱癌 (bladder cancer)			

续表

研究对象和目的	材料	方法与结果	参考文献
建立膀胱癌患者尿液中的蛋白质数据库	约50例正常人和膀胱癌患者的尿样	样品透析冻干，2-DE、微量测序和MALDI-TOFMS鉴定找到有表达差异的蛋白质包括psoriasin，psoriasis-associated acid-binding protein，gelsolin，prostaglandin D_2 synthetase	[93]
建立在变移性细胞癌(TCC)和鳞状细胞癌(SCC)中上调和下调的角蛋白质数据库	变移性细胞癌(95%在西方国家)，鳞状细胞癌(80%在非洲和中东)	2-DE、微量测序和MALDI-TOFMS鉴定和免疫组化分析	[94,95]
建立一个诊断过渡性细胞癌的试验	共94例正常人、变移性细胞癌的膀胱癌患者和其他泌尿生殖病患者的尿样	SELDI-TOFMS（使用强阴离子交换芯片）找到5个蛋白质和5个蛋白质簇偏好在变移性细胞癌细胞中表达；使用不同组合的这5种蛋白质和蛋白质簇，可以高灵敏度地检测到早期的变移性细胞癌	[96]
前列腺癌(prostate cancer)			
寻找前列腺癌的分子标记	正常前列腺组织和前列腺癌组织	LCM，SELDI-TOFMS检测PSMA和几个10×10^3Da以下的蛋白质只在癌细胞中表达，33 436Da的蛋白质在癌组织中表达量上调	[97]
寻找前列腺癌的分子标记	正常前列腺组织、恶性前的前列腺上皮细胞瘤(premalignant prostatic intraepithelial neoplasia，PIN)和具有侵袭力的前列腺癌组织	LCM，SELDI-TOFMS发现两个约28和32×10^3Da的蛋白质在正常和癌组织中有差异表达，在PIN细胞里它们的表达呈过渡态	[98]
雄激素处理对前列腺癌细胞的影响	雄激素饥饿和刺激的前列腺癌细胞系M12AR	2-DE，ESI-MS/MS和SEQUEST数据库搜索鉴定出雄激素诱导表达的一种蛋白质human nucleoside diphosphate kinase A（NDKA)，可抑制肿瘤转移	[99]
胃肠道肿瘤(gastrointestinal malignancies)			
食道癌(esophageal cancer)			
寻找与恶性肿瘤生长相关的蛋白质	正常鳞片状上皮细胞及相应的二期肿瘤细胞	LCM，2-DE，ESI-ITMS找到7个蛋白质仅在正常细胞中表达，10个蛋白质在肿瘤细胞中有显著差异，两个鉴定出为cytokeratin1和annexin I；LCM，SELDI-TOFMS比较正常细胞和肿瘤组织，找到2个蛋白质在肿瘤中表达量上调	[99,100]
结肠癌(colon cancer)			
LCM与标准切片技术相结合的效用，寻找肿瘤恶性生长相关的蛋白质	正常结肠上皮细胞和肿瘤细胞	LCM与2-DE，MALDI-TOF联合应用，不染色，可相对减少非上皮的污染蛋白(如血红蛋白)，同时增加细胞内上皮结构蛋白如胞浆角蛋白的显示	[101]
建立人结肠癌细胞膜蛋白质数据库	人结肠癌细胞系LM1215	SDS-PAGE，HPLC和ESI-ITMS. 鉴定了92个膜蛋白. 使用“target ion”技术可鉴定到低丰度蛋白质，如A33抗原	[102]

续表

研究对象和目的	材料	方法与结果	参考文献
正常和癌变的结肠黏膜细胞的比较，寻找与肿瘤恶性生长相关的蛋白质	正常和癌变的结肠黏膜细胞	2-DE，MALDI-TOFMS 鉴定到在癌细胞中高表达的 S100A9，S100A8 和 HSP70 的一种新变异体。癌细胞中表达下调的蛋白质包括肝脂肪酸结合蛋白，肌动蛋白结合蛋白等	[103,104]
由 APC 和 p53 突变引起的结肠上皮细胞的 crypts and 息肉的蛋白质变化	正常、p53 缺失，和多重结肠瘤(multiple intestinal neoplasia，MIN)的老鼠	2-DE，LC-ESI-ITMS 发现患癌动物的息肉中有约 60 种蛋白质显著下调，4 种蛋白质显著上调，鉴定了 38 个蛋白质 carbonic anhydrase I 在正常 crypts 中显著下调	[105]
与息肉形成相关的蛋白质	正常结肠上皮细胞和人息肉	2-DE，MALDI-TOFMS：在息肉中表达上调的包括 numatrin，HSP70 和 HSP60；下调的蛋白质包括脂肪酸结合蛋白 14-3-3ä，cytokeratin 20，Cyt. c 氧化酶等 6 种	[106]
与细胞化学抗性相关的表达下调的蛋白质	直肠肉瘤细胞系 HT-29RNOV，对 mitoxantrone 有非典型抗性	2-DE，MALDI-TOFMS 鉴定了一个在抗性细胞中下调的蛋白质 human adenine phosphoribosyl transferase	[107]
	直肠肉瘤细胞系 HT-29DRB，对 daunorubicin 有抗性	2-DE，MALDI-TOFMS 鉴定出两个在抗性细胞中下调的蛋白质，一个是人乳腺癌特有基因 1 的同源蛋白，另外一个为酵母蛋白 yer-1 的同源蛋白	[108]
用 SELDI-TOF MS 在血清寻找肿瘤分子标记	51 例良性、恶性和前恶性的结肠瘤患者的血清	用结合阴离子蛋白质的 SAX2 类型芯片的 SELDI-TOFMS 找到一个 13.8×10^3Da 的蛋白质在结肠癌和息肉以及憩室炎细胞中均表达，而良性疾病和正常对照细胞中不表达	[109]
肝癌 (liver cancer)			
肝癌细胞系的蛋白质组数据库	HCC-M	2-DE、MALDI-TOFMS 鉴定了 272 个银染的蛋白质点，用 nESI-MS/MS 和 *de novo* 测序鉴定了剩余的全部 29 种蛋白质。发现一种新的核蛋白 HCC-1，其表达与肝细胞癌的分化程度呈负相关	[110]
与肝癌恶性生长相关的蛋白质	人肝癌细胞系 BEL-7404 和永生化的正常肝细胞系 L-02	由 LC/ESI-ITMS 获得 CID 数据，搜库鉴定。肿瘤细胞中 GST P 表达量上升 18 倍，另有 6 种表达上调，3 种下调；2 种脂肪酸结合蛋白只在正常肝细胞中表达	[111]
胰腺癌 (pancreatic cancer)			
和胰腺癌转移和复发相关的蛋白质	胰腺癌的转移和复发通常和产生化学抗性相关，这里所用的是人胰腺腺瘤细胞培养模型 EPP85-181P，对 daunomycin 有抗性和对 mitoxantrone 有抗性的两个亚细胞系	2-DE，MALDI-TOFMS 发现 cofilin 在两个亚细胞系中均过量表达，另外有 2 种蛋白质只在对 mitoxantrone 有抗性的细胞中表达	[112]
其他肿瘤			

续表

研究对象和目的	材料	方法与结果	参考文献
肺癌 (lung cancer)			
与肺癌发生相关的分子标记	正常肺和肾组织与原发性的肺腺瘤组织	2-DE、ESI-ITMS显示TAO1在大于90%的原发性肺腺瘤中表达,但在其他恶性肿瘤中不表达。TAO1/2可能用于鉴别诊断原发性和转移性的肺腺瘤	[113]
寻找肺癌发生的分子标记	手术切除的正常肺组织和肺癌组织	2-DE、MALDI-TOFMS和ESI-ITMS发现Cytokeratin 8和Cathepsin D在肺腺瘤中表达量上调并有变异,认为后者促进了肿瘤的侵袭和转移。在小细胞瘤里,异源型核糖蛋白HnRNP表达量升高了3倍;在非小细胞瘤系A549中检测到一种在乳腺癌中表达的α-tubulin isotype	[86,89]
肺上皮细胞对辐射产生的保护反应的机制	正常支气管上皮细胞和受0.5 Gy辐射的上皮细胞(L132细胞)	2-DE-MALDI-TOF找到至少3种表达上调的蛋白质与PKC信号通路相关,其中蛋白激酶抑制剂1起主要作用,另有3种表达下调的蛋白质	[114]
恶性黑素瘤 (melanoma)			
黑素瘤发生的分子标记	LCM微切割下的黑素瘤和其周围正常上皮部分	通过SELDI-MALDI-TOFMS(应用反相芯片或强阴离子交换表面芯片)获得蛋白质表达谱,检测到4个可能的分子标记,但这些分子的意义和显著性尚未分析	[115]
黑素瘤发生的血清分子标记	正常对照人的血清和黑素瘤患者的血清	血清用分子排阻的超滤方法来分级,MALDI-TOFMS分析找到7个小于35×10^3Da的蛋白质在病人血清中特异性表达;某些2.5×10^3~3.5×10^3Da的多肽和疾病的发展有显著变化,但没有做鉴定	[116]
与转移的恶性黑素瘤的化学抗性相关的蛋白质	从黑素瘤患者淋巴中建立起来的细胞系和4种分别对vindesine, cisplatin, fotemustine, 和etoposide有抗性作用的亚细胞系	2-DE、考马斯亮蓝染色、MALDI-TOFMS发现至少7种蛋白质在对照细胞中未找到,有4种在抗性细胞系中过表达	[117]
头颈部肿瘤 (head and head cancer)			
和肿瘤发生相关的蛋白质	粗的肿瘤组织和正常组织	样品未经LCM纯化,SELDI-MALDI-TOFMS分析,所以重复性较差,一个8.7×10^3Da的蛋白质在5/6例肿瘤组织中,6例正常组织中未见到	[115]
神经胶质母细胞瘤(glioblastoma)	人神经胶质母细胞瘤细胞系移植到裸鼠中,肿瘤长成1cm直径大小时,肿瘤被切成12μm厚的切片,肿瘤增殖的中心部位和边缘部位,以及缺血性坏死区	肿瘤增殖的中心部位和边缘部位,以及缺血性坏死部位的切片通过MALDI-TOFMS做MS图谱,鉴定到超过150种蛋白质;对感兴趣的蛋白质再做ESI-ITMS,thymosin α4在肿瘤增殖的中心区高表达	[81]
成神经细胞瘤 (neuroblastoma)			

续表

研究对象和目的	材料	方法与结果	参考文献
ProNPY 的加工和成神经细胞瘤发展不同阶段的关系	成神经细胞瘤组织和良性对照以及神经节细胞瘤(ganglioneuroma)组织	分离具有类似 NPY 免疫活性的肽段并用 MALDI-TOF 鉴定,找到其中一个为 NPY3-36,其受体选择性不同于完整的 NPY。NPY 的前体 proNPY 的加工受到抑制与肿瘤的转移发展相关	[118]

本实验室从 1998 年开始对肝癌恶性生长相关的的表达蛋白质组展开研究,以人肝癌细胞系 BEL-7404 和正常肝细胞系 L-02,以及恶性生长特性部分逆转的 JX-1 细胞株与其对照细胞株 JX-0 为实验材料[111,119]。JX-1 是 EGFR cDNA 序列以反义方向转染 BEL-7404 人肝癌细胞而建立的细胞株。通过双向凝胶电泳和数字化图谱的软件分析,发现 7 个蛋白质斑点只在 BEL-7404 细胞中检测到表达,14 个斑点只在 L-02 肝细胞中检测到表达,78 个斑点在两细胞系间有显著的表达量差别,并由此建立了蛋白质组的差异表达谱。应用胶内酶解和质谱(MS)分析技术测定蛋白质斑点的 MS/MS 数据,结合数据库搜索,12 个蛋白质斑点从中得到鉴定。进一步通过与 JX-0 细胞对比,发现 40 个蛋白质斑点在 JX-1 细胞蛋白质组中发生明显的表达变化,其中 3 个斑点的表达趋于正常化。凝胶配比和质谱鉴定揭示肿瘤抑制蛋白(maspin)在肝癌细胞中表达量显著下降,而在恶性生长特性部分逆转的细胞中又有较大回升;凝胶上有两个斑点为 HSP27,它们在肝癌细胞的量高于肝细胞或逆转细胞,HSP27 的变化与它不同的化学修饰有关,显示蛋白质翻译后修饰的动态变化对癌细胞生长的潜在影响。我们首次发现 maspin 和 reticulocalbin 同源蛋白在肝细胞表达和在肝癌细胞中的表达变化,以及 maspin 在肝癌生长特性部分逆转的细胞中趋于正常化的表达。对肝癌细胞和肝癌恶性生长特性部分逆转的细胞蛋白质组变化的分析有助于肝癌发生机制的阐明,并为建立相应的蛋白质组数据库打下了基础。目前正在开展与功能关系密切的肝癌细胞模型和临床肝癌组织样品的蛋白质组研究,以寻找与肝癌发生和预后相关的分子标记。

随着蛋白质组学技术在肿瘤研究中的应用,新的肿瘤分子标记将不断被发现,为了更好地用于临床,有必要建立一些标准对这些分子进行筛选和确认。美国国立肿瘤所(national cancer institute)领导建立了早期检测研究网络(early detection research network)的组织,他们提出 5 条标准来发展和评估肿瘤分子标记:①临床前的调查阶段,找出合适的研究方向;②临床分析和评估阶段,看看临床分析是否能检测已知的疾病;③纵向的回顾阶段,看看假定的肿瘤分子标记是否能检测临床前的疾病并且建立阳性筛选准则;④通过试验和相关的阳性/假阳性率来建立预期的筛选,确认疾病发生的性质和程度;⑤设计最终试验(随机试验)看看该筛选方法对降低普通人群中某种疾病的发生到底有多大效果[120]。

同时鉴于蛋白质组图谱诊断在早期肿瘤检测方面的显著临床潜力,美国国立癌症研究所将于 2004 年开展对卵巢癌和其他肿瘤的临床试验,以评估该技术的实用性。而目前一些主要的实验室(reference lab)已经将该技术运用到常规检测中[121]。

第三节　蛋白质组学在神经系统研究中的应用

一、介　　绍

大规模工业化基因测序的成功为生物系统的研究开辟了新的道路。在后基因组时代,

我们应当研究 DNA 序列是如何翻译为细胞、组织和个体生长发育所需的信息。其中的一个研究重点就是确定蛋白质功能及其分子网络。新兴的蛋白质组学技术与数据库挖掘的结合大大加快了这一研究进程。神经系统由于其高度的复杂性和同一突触存在不同神经递质之间的相互作用，尤其适合用蛋白质组学等方法来研究。目前已知的数据库包括果蝇脑、大鼠小脑蛋白质组数据库(http://www-lecb.ncifcrf.gov)，小鼠少突神经、视神经蛋白质组数据库及新生小鼠脑、小脑蛋白质组数据库等(http://proteomics.cancer.dk)。

与传统方法相比，现代蛋白质组学技术的最大区别在于它的高通量，能够同时鉴定出的蛋白质种类数量巨大。关于神经活动的研究，如学习过程的研究，早在 19 世纪 50 年代就开展了，它通过玻璃毛细管电泳进行[122,123]，结果表明有新蛋白质的合成，但只观察到这些蛋白质含量的变化，并且这些蛋白质的种类未知。19 世纪 80 年代，几个实验室通过分级分离、生化和免疫化学等方法[124~132]在多种有机体中鉴定了突触蛋白质。这应当归功于分离技术的提高，从而一些重要的区室结构，如突触小泡、突触后致密体等都能够被分离。特异蛋白质的鉴定增加了突触组分的数目。目前，突触小泡的蛋白质组分已经鉴定出来[133]。在突触后区域，Edman 测序或质谱方法鉴定出突触后致密体中大约 30 种蛋白质[134]。

二、应用举例

蛋白质组学技术方法的出现给神经科学的发展带来新的动力。在神经系统中，很多功能特异而复杂的结构为蛋白质组学分析提供了良好的材料，例如，在近郎飞结处轴突胶质细胞接头适合于作蛋白质组学分析[135]。

(一) 对 NMDA 受体多复合物(NMDA receptor multiprotein complexes，NRC)的研究

蛋白质组学可用来研究与受体相关的信号复合体。在大脑中，以受体为核心由多个蛋白质组成的信号复合体在信息处理过程中起着重要作用[136]。一个已知的例子是 NMDA 受体与突触后蛋白质结合，其中包括 PSD-95。该复合体与突触可塑性和学习有关。Husi 等人在研究小鼠的前脑部位时，利用蛋白质组学方法获得了 NRC[137]。复合体的蛋白质之间发生相互作用，有时需要翻译后修饰，如磷酸化(PI3 激酶的 SH2 结构域结合到 NMDAR 亚基上[138])、脂类修饰[139]等。用蛋白质组学的方法可以检测到此类相互作用以及双杂交筛选检测不到的微弱相互作用[140]。

目前，蛋白质组学包括了多种大规模鉴定蛋白质的方法。Husi 等人比较了从小鼠脑中分离纯化 NRC 的不同方法[137]，包括：标准的 HPLC 色谱、含 NMDAR 配体的亲和色谱、含纯化抗体的免疫亲和色谱、大规模的免疫共沉淀、肽段亲和纯化。其中，以下三种方法检测效果较好：免疫亲和色谱、以抗 NMDA 受体亚基抗体进行的免疫沉淀、基于 NMDA 受体 2B 亚基 C 末端的肽段亲和纯化[141]。SDS-PAGE 分离各组分后发现复合体的蛋白质组成复杂。NRC 是一个相对分子质量约为 2×10^6Da 的复合体[137,141]，由 75 种或更多的蛋白质组成，分为神经递质受体、细胞黏连分子、接头蛋白、信号通路成分、细胞骨架蛋白及一些其他蛋白质[142]。复合体中，至少 27 种蛋白质已知与突触可塑性相关，而另有 18 种是学习过程所必需的。通过对复合体中蛋白质的鉴定，Grant 等人提出[142]：①复合体的组成提示多蛋白质复

合体能够检测神经细胞活动的不同模式，并且能与相应的信号传导通路相连，改变神经元的状态；②在学习和突触可塑性的生物学机制中，复合物本身是一个重要的组成部分；③对上述任何组成成分的干扰都将影响到复合体的整体功能。

国外学者研究发现 NRC 是神经元一个特殊的输入系统[142]，它能检测并区分神经元活动的不同模式，然后激活不同的下游信号通路(图 23.2)，改变神经元特性，储存活动模式中编码的信息。神经元的变化包括一系列不同的过程，这其中有突触本身的变化(突触可塑性)、树突局部的变化和整个神经元的变化(如基因转录)。由 NRC 中信号分子的多样性分析以及遗传学与药理学干扰实验结果表明，NRC 的一个重要特性就是可以将神经递质受体获得输入信号后所引发的不同下游信号通路的变化协调起来。

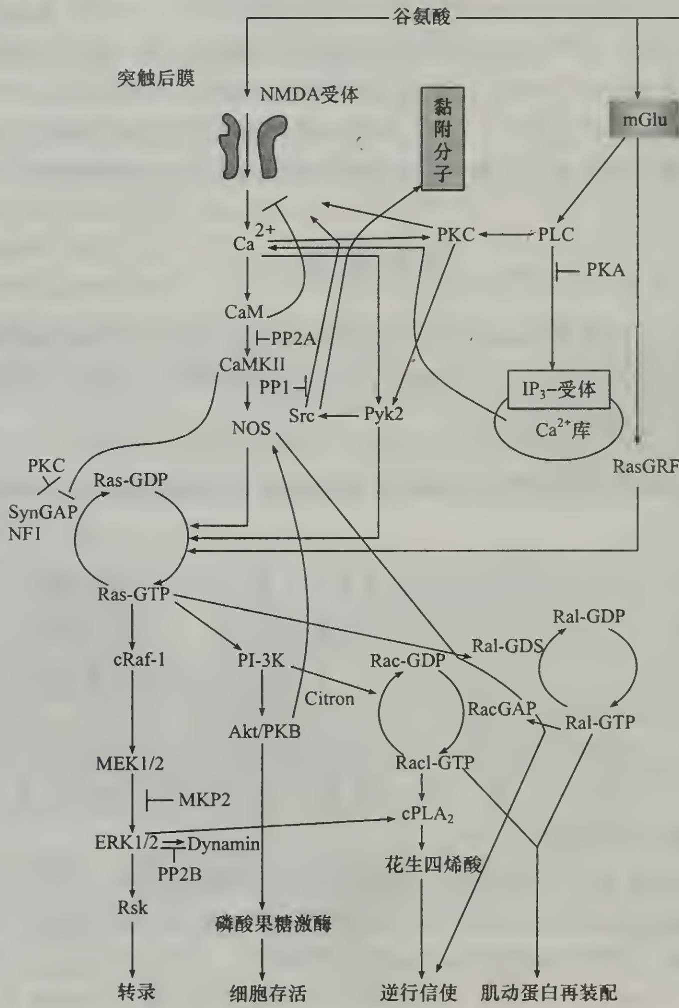


图 23.2　NMDA 受体复合物介导的信号通路[143]

图中的浅色框表示 NMDA 受体复合物中的蛋白质。在 NMDA 受体复合物中有许多已鉴定出的蛋白质参与了 NMDA 受体下游的信号通路。CaM，钙调蛋白；CaMKⅡ，钙-钙调蛋白激酶Ⅱ；ERK，胞外信号相关激酶；IP3，1，4，5-三磷酸肌醇；MAP，微管相关蛋白；MEK，MAPK/ERK 激酶；mGlu R，代谢型谷氨酸受体；nNOS，一氧化氮合成酶；PLA，磷脂酶 A；PLC，磷脂酶 C；PKA，蛋白激酶 A；PKC，蛋白激酶 C；PP1/2A/2B，蛋白磷酸酶 1/2A/2B

（二）对海马区的研究

双向凝胶电泳技术(two-dimensional electrophoresis,2-DE)能对细胞或组织中表达的所有蛋白质进行整体分析。对于这项技术人们所关注的是电泳图上最多能分辨出的蛋白质点的个数以及怎样去检测低丰度的蛋白质。一张标准的大鼠海马神经元的 2-DE 图通过银染后能分辩出 853 个点[144]。为了提高分辨率和低丰度蛋白质的检测效果,Takashi 等人运用改进的 2-DE 方法,第一向使用 7 个不同的狭窄 pH 范围(pH 3.5～4.5,4.0～5.0,4.5～5.5,5.0～6.0,5.5～6.7,6～9,6～11),第二向使用 14 块 SDS-聚丙烯酰胺凝胶。7 块 7.5%的胶用于分离分子质量较高的蛋白质($\geq 40\times10^3$Da),7 块 13.5%的胶用于分离分子质量较低蛋白质($<40\times10^3$Da)。14 块胶的图像拼成一块 cybergel(70cm×67cm),包含 6 677 个蛋白质点,与标准的 2-DE 图谱相比,改进后的 2-DE 检测到的蛋白质点数目增加了 783%[144]。

在低氧条件下,海马 CA1 和 CA3 区的反应是不同的。CA1 更为敏感,尤其是在间歇性低氧情况下。Gozal 等人分析了大鼠海马 CA1 和 CA3 区差异表达的蛋白质,并比较了在进行 6h 间歇性低氧处理后蛋白质表达水平的变化。分析后发现在含氧量正常的情况下,99 个蛋白质同时在 CA1 和 CA3 区表达。在两区中具有表达差异性的蛋白质归为与细胞骨架相关的蛋白质和与代谢相关的蛋白质两类(表 23.3)。这两类蛋白质在 CA1 区的表达量较高。6h 间歇性低氧处理后,CA1 中 32 种蛋白质表达升高,而 CA3 中只有 7 种蛋白质的表达量升高[145]。

表 23.3　含氧量正常情况下,在 CA1 和 CA3 差异性表达的蛋白质[145]

在 CA1 和 CA3 区差异性表达的蛋白质	CA1/CA3 密度测定	P
与细胞骨架相关的蛋白质		
TOAD 64	0.4	0.03
Alpha tubulin 1	4.3	0.01
Beta tubulin 1	5.3	0.01
Myosin short subfragment 2	1.9	0.01
Glial fibrillary acidic protein	0.4	0.03
Fascin homolog 1	1.7	0.04
Leukemia assoc gene (Stathmin)	5.7	0.03
与代谢相关的蛋白质		
Guanine deaminase	4.0	0.009
Fructose bisphosphate aldolase c1	0.2	0.004
Pyruvate dehydrogenase E1 component	1.9	0.05
Phosphoglycerate mutase 1, brain	2.3	0.04
Triose phosphate isomerase	3.8	0.03
Cytosolic malate dehydrogenase	2.3	0.04
Alpha enolase	2.8	0.01

Edgar 等人运用比较蛋白质组学的方法分析患有精神分裂症个体的海马蛋白质组后发现,精神分裂症与 6 号染色体长臂有关[146]。他们比较了 7 个患病个体和 7 个对照个体,发现在患病个体的海马区,18 个蛋白质的浓度有显著变化($P<0.05$)。通过 N 端测

序，鉴定了其中三种蛋白质与 6 号染色体长臂相关：安定结合抑制剂，基因位置 6q12-q21；锰超氧化物歧化酶，基因位置 6q25.3；T-复合物蛋白 1，基因位置 6q25.3-q26。

（三）对唐氏综合征的研究

唐氏综合征(Down syndrome，DS)，即 21 号染色体三体，是智力发育阻滞最普遍的遗传因素。通过分析脑神经病理学、神经学、神经生理学、神经药理等变化，发现 DS 患者的中枢神经系统的结构发育和功能异常[147~151]。虽然在 DS 成人脑中已经发现一系列的因子和蛋白质表达异常，但是 DS 婴儿脑的蛋白质组学信息却非常有限。

Cheon 等人在蛋白质水平上采用“基因追踪”(gene hunting)的方法，建立了脑蛋白质图谱，定性定量分析了双向凝胶电泳图谱上不同的蛋白质点(图 23.3)[152]。鉴定分析出

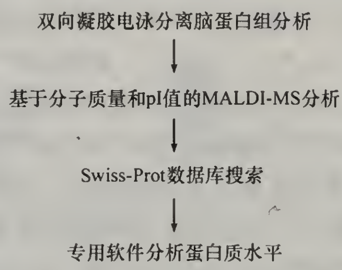


图 23.3　分析 DS 婴儿的蛋白质组学方法[152]

的 10 个蛋白质中 septin6 和 DKFKZp564D177.1(片段)表达异常[152]。DKFZp567D177.1(片段)的水平在 DS 婴儿脑中大大降低，而脑中其他 septin 蛋白质的水平与对照组差别不大(表 23.4)。septin 6 在 DS 婴儿脑中的出现率低，对照组中 7 个正常婴儿脑出现率为 100%，而在患病组仅为 1/7。septin 是一类进化上保守的细胞骨架 GTP 酶，与胞质分裂、细胞极性、胞吐作用相关[153]。不同的 septin 显示出细胞骨架和膜联系的不同模式[154]，比较免疫组化分析表明，该家族不同成员在脑中起着不同的作用[155]。

表 23.4　DS 婴儿脑的蛋白质水平[152]

蛋白质	对照(个数)	DS(个数)
WD 重复蛋白 1	3.03+1.66(7)	3.54+1.32(6)
与 septin 同源的新蛋白质	2.08+0.69(7)	2.39+0.70(5)
septin 5	10.74+0.47(7)	11.56+0.35(5)
septin 2	1.07+0.63(7)	1.54+0.48(4)
DJ37E16.5	1.81+0.35(7)	1.64+0.38(5)
30.2×10^3Da 的假设蛋白质	1.49+1.18(7)	0.79+1.22(8)
DKFZp564D177.1(片段)	3.43+1.83(7)	0.61+0.44*(7)
神经蛋白 NP25	1.62+0.54(7)	1.39+0.76(6)
DC7 蛋白(囊泡分拣蛋白 29)	1.27+0.64(7)	1.16+0.39(8)

双向凝胶电泳分离 DS 婴儿和正常个体的大脑皮层中的蛋白质，然后经考马斯亮蓝染色。在部分胶图像上，用 ImageMaster 2D 软件定量分析蛋白质点的量。用 Graphpad Instat2 软件进行数据处理。(*)号表示存在显著性差异($P<0.01$)。

Septin 6 与不同脑结构中突触小泡的形成有关，尤其是海马结构 CA1，CA2[155]，而 DS 患者的显著标志是突触结构异常[156]，由此我们可以得出结论：Septin 6 蛋白质水平的变化一方面引起了脑结构的变化，另一方面又可作为 DS 的诊断标志。

(四) 对阿尔茨海默病的研究

阿尔茨海默病(Alzheimer disease，AD)又名进行性老年性痴呆，在所有痴呆症中占大约 50%，是最常见的一种痴呆性病症，危及到约 400 万名老年人的健康。AD 患者在出现明显认知障碍和其他临床症状之前，大脑中就存在异常改变。但是目前对它的发病机制却了解甚少。

活性氧或其他活性物质对蛋白质的氧化修饰会引起蛋白质功能丧失或异常，是衰老或衰老相关的神经衰退性疾病的可能原因之一[157~161]。蛋白质羰基化的形成可作为蛋白质氧化的检测标志[162]。AD 患者中，羰基化程度升高[162,163]。Castegna 等人[164]将 2-DE 印迹、羰基的免疫检测技术及质谱技术相结合，鉴定了在 AD 中特异性氧化的蛋白质(表 23.5)：肌酸激酶 BB(creatine kinase BB，CK BB)，谷氨酰胺合成酶(glutamine synthase，GS)，泛素 C 端水解酶(ubiquitin C-terminal hydrolase，UCH L-1)。其中，CK BB 的羰基化程度增高 4.9 倍。CK BB 与 ATP 的合成相关。AD 患者中，CK BB 氧化后活性下降，说明在氧化胁迫下，可供神经元和突触元件使用的能量减少。CK BB、GS、UCH L-1 是细胞代谢中重要的蛋白酶，自由基对这三个蛋白质的修饰是 AD 病的一个临床症状，并且是引起神经衰退的机制之一。

表 23.5　与对照相比，AD 患者脑中三种蛋白质羰基化水平的相对改变[164]

鉴别的蛋白质	特异性的氧化	P 值
肌酸激酶(CK BB)	489±52	<0.01
谷氨酰胺合成酶(GS)	321±18	<0.01
泛素 C 端水解酶(UCH L-1)	224±63	<0.02

在啮齿类动物中，一些神经类固醇已经显示有神经保护的作用[165,166]。为了解 AD 病中神经类固醇的生理重要性，Weill-Engerer 等人[167]采用 GC-MS 比较了 AD 患者和对照个体不同脑区域(海马、杏仁核、额皮质、纹状体、下丘脑和小脑)的 6 种类固醇的量：孕烯醇酮(pregnenolone)、孕烯醇酮硫(pregnenolone sulfate，PREGS)、脱氢异雄酮(dehydroepiandrosterone)、脱氢异雄酮硫(dehydroepiandrosterone sulfate，DHEAS)、孕酮、异孕烯醇酮。与对照相比，AD 患者的所有神经类固醇普遍降低：PREGS 和 DHEAS 在纹状体和小脑中的含量显著降低，下丘脑中的 DHEAS 含量也显著下降。纹状体和小脑中皮质 β-amyloid 肽段和 PREGS 含量水平呈现显著的负相关。同时下丘脑中磷酸化的 tau 蛋白和 DHEAS 也呈现负相关。与神经纤维缠结(老年性痴呆病的一种临床症状)相关的特征蛋白质的高水平的表达，往往伴随着脑中 PREGS 和 DHEAS 水平的降低，说明在 AD 中神经类固醇也许起着神经保护的作用。

为了更清楚地了解 AD 的致病机制，Schonberger 等人[168]分析了 6 种脑组织区域蛋白质组学方面的变化。他们选取了 AD 患者死后的脑组织和相同年龄段的正常人脑组

织，采用定量蛋白质组学方法比较个体蛋白质的信号变化。在严重受损的脑区域，与对照组相比较，76 种蛋白质在 AD 患者海马结构中特异性表达，62 种蛋白质在颞皮质特异性表达，39 种蛋白质在内嗅区皮质特异性表达。在相对稀疏的区域也存在着显著差异，与对照相比，34 种蛋白质在小脑中特异表达，125 种蛋白质在扣带回区特异性表达，75 种蛋白质在感觉运动区皮质特异性表达。鉴定了其中 37 种蛋白质，分析了关键致病途径变化的相关性。

（五）对帕金森病的研究

帕金森病（Parkinson's disease，PD）是中老年人的一种常见病，临床表现以静止性震颤、面部表情僵化、肌强直、姿势反射减少、呈进行性加重为主要特点。PD 的一个生理指标是黑质和被盖中产多巴胺的细胞衰退，导致多巴胺能投射区的多巴胺不足，特定皮质-亚皮质锥体束外环路的功能下降[169]。黑色素及神经细胞减少达 59%，神经胶质细胞增生，在黑质和兰斑等细胞质内发现包含体（lewy body），约占 90%。

自 1983 年 Langston 等人[170]发现接触含有 1-甲基-4-苯基-1,2,3,4-四氢吡啶（MPTP）的物质可引起 PD 后，经过多年的研究，MPTP 的神经毒性已经明确，可以选择性损害黑质、纹状体系统，导致多巴胺缺乏。Kajita 等人通过气相色谱-质谱（gas chromatography-mass spectrometry，GC-MS）技术，首次在 PD 患者脑和正常脑中分离到 *N*-甲基多巴胺，而 MPTP 的同源物 *N*-甲基猪毛菜酚和 *N*-甲基去甲猪毛菜酚，可能由 *N*-甲基多巴胺通过 Pictet-Spengler 凝缩反应得到[171]。

液相色谱与多电极检测器联用，可用于定量分析 PD 患者脑室液中神经化学物质。Ikarashi 等人研究发现，在高效液相色谱图谱中，有一未知峰与 5-羟色胺有相同的保留时间。通过离子抑制（ion suppression）、快原子轰击质谱（fast atom bombardment mass spectrometry）分析后，鉴定该成分为 *O*-甲基多巴，为 *L*-多巴的主要代谢产物[172]。

Zhou 等人[173]采用 GC-MS 分析 PD 患者脑脊液中 β-苯乙胺（PEA）的量，评价脑脊液中 PEA 浓度与 PD 严重性的相关程度。共分析了 23 名患者（平均年龄 64.0，标准差 8.2）的脑脊液，3 名患者处于 Hoehn 和 Yahr 阶段Ⅱ，11 名患者处于阶段Ⅲ，其余 9 名患者处于阶段Ⅳ。同时比较了 8 名神经病患者（平均年龄 57.0，标准差 19.2）和 12 名正常个体（平均年龄 57.6，标准差 4.8）脑脊液中 PEA 的含量。结果发现，PD 患者脑脊液中的 PEA 浓度（205pg/ml）显著低于神经病患者（433pg/ml）和正常个体（387pg/ml）。处于 Hoehn 和 Yahr 阶段的 PD 患者的 PEA 浓度呈现出负相关（$P<0.01$）。Miura 等人[174]在研究 PD 患者胞浆 PEA 含量时也发现，在病情加重的 PD 患者中，PEA 的含量呈现出下降的趋势。而在病情未加重的患者中，PEA 含量维持在同一水平。这一结果表明 PEA 水平的下降可能反映了 PD 患者黑质纹状体的衰退。

第四节　蛋白质组学在体液研究中的应用

一、介　　绍

（一）体液的基本知识

机体含有大量的水分，这些水和溶解在其中的各种无机物、有机物及电解质、非电解

质总称为体液(body fluid),约占体重的60%。体液以细胞膜为界,可分为两大部分,即细胞内液(intracelluler fluid,ICF)和细胞外液(extracelluler fluid,ECF)。存在于细胞内的称为细胞内液,约占体重的40%。存在于细胞外的称为细胞外液。

细胞外液因其存在部分不同,又可分为两类:一类是存在于组织细胞之间的组织间液(包括淋巴液和脑脊液),约占体重的16%。细胞间液是血浆的超滤液,其电解质成分和浓度与血浆极为相似,不同之处是血浆含有较多的蛋白质,而细胞间液不含或仅含少量的蛋白质。由于蛋白质是相对分子质量较大的物质,不易通过细胞膜,故血浆蛋白含量高于细胞间液。另一类是血液的血浆(约占体重的4%)。血液是结缔组织的一种,由无形成分和有形成分,即占血液总量的55%血浆和占血液总量的45%血细胞两部分组成。血浆中溶解有多种化学物质,按容积计算,其中包括:水(91%)、蛋白质(7%)、脂质(1%)、糖类(0.1%)、非蛋白氮NPN(1%)、无机盐类(0.9%)和代谢产物(尿素、肌酐、尿酸等)。其中血浆蛋白有白蛋白、球蛋白、纤维蛋白原三类。白蛋白在维持血浆胶体渗透压中有重要作用,球蛋白与抗体产生有关,纤维蛋白原参与血液凝固。血浆与血清的成分差别仅在于后者因采血后未加抗凝剂而不含纤维蛋白原。

(二)体液与生物医学

正常状态下,人体内由外界摄入的水分及电解质与体内氧化产生的水(内生水),不断与各区间的体液进行交换。体液交换包括血浆与细胞间液、细胞间液与细胞内液之间的交换,前者交换的动力是血浆胶体渗透压与静水压(血压)之差,其中胶体渗透压在血浆与细胞间液的交换中起主要作用,同时还可影响细胞外液的总量。细胞间液与细胞内液之间的交换主要决定于细胞内外液的渗透压。正常体液的分布、组成及容量三方面均在神经体液等因素调节下保持动态平衡,以保证机体各种生理活动正常进行。体液平衡是维持机体生命活动的必不可缺少的条件,疾病发生的基本机制之一就是体液机制,即病因引起体液质和量的变化、体液调节障碍,最后造成内环境紊乱而致病。我国传统的中医理论也指出:人分阴阳,阳指身体的机能,阴则指体内的体液,包括血液、唾液、泪液、精液、内分泌液等。中医学以阴阳鱼为思维模式,阐述人体生理或病理变化。人的健康状态是阴阳处于动态平衡的历调和谐状态,即人体内环境的稳定状态;阴阳失调则是人体的疾病状态,常表现为阴盛、阳盛、阴虚、阳虚四种类型。

已有的医学理论与迄今为止的疾病研究诊治实践都表明,体液与人的生理、病理都有着密切的关系。对各种体液质与量的全面深入的了解是人类理解自身生命活动规律、追踪生理或病理变化的突破口之一。而建立并注释各种体液的全套蛋白质组学图谱,既推动了生物医学基本理论研究的进程,又为临床领域提供了多种可能的诊断标志物、预后评估指标和药物作用靶点,有着极为深远的理论与实践意义。

(三)应用蛋白质组学的方法研究体液

蛋白质组是指基因组表达的所有相应的蛋白质,是指细胞或组织或机体全部蛋白质的存在及其活动方式。蛋白质组学是从整体的蛋白质水平上、从生命本质的层次上,研究和发现生命活动的规律和重要生理、病理现象的本质。这与疾病发生的体液机制及中医学的阴阳鱼思维模式可谓是不谋而合。现代蛋白质组学的发展更为体液研究提供了越来

越多的技术平台[175~178]。

蛋白质组学的研究不局限任何特定的方法。高分辨率的蛋白质分离技术如二维凝胶电泳和高效液相层析，经典的蛋白质鉴定方法如氨基酸序列分析等，现代质谱技术，基因组学研究的各种手段，现代计算机信息学和计算机网络通讯技术，等等，任何可用于蛋白质研究的技术手段，蛋白质组学都可能采用[175~178]。它体现的是一个开放的思维和研究方式，而体液的多样性和复杂性则为各种蛋白质组学技术提供了无限广阔的应用土壤。

二、体液的蛋白质组学研究现状

今天，世界各地的大学、研究机构和公司都将目光集中在这一领域，涉及血浆、血清、脑脊液、尿液、唾液等几乎所有种类体液的蛋白质组学研究正蓬勃发展起来。由于细胞内液的不易分离，并且它的功能与细胞整体密不可分，因而体液的蛋白质组学研究重点在血浆与各种组织间液。另外，临床使用的血清因其广泛的应用价值，也被归于体液的蛋白质组学的研究内容。在此我们简介各种体液蛋白质组学的研究现状。

（一）有关血浆和血清

蛋白质组学的形成已有近 10 年的历史，就像生命科学的其他分支一样，蛋白质组学的最终目标都是为了更深入的对生命进行认识，从而能为人类自身的健康事业提供支持。蛋白质组学的研究样本多种多样，无一例外都是生物样本。在蛋白质组自身的发展中，不断地追寻着最能反映生命本质的生物样本。血液通过循环系统与全身各个组织、器官密切联系，参与机体各项生理活动，维持机体正常新陈代谢和内外环境平衡。在某些病理情况下，血液系统疾病除直接累及血液外，还可影响全身组织、器官，而各组织、器官的病变也可直接或间接地引起血液发生变化。今天，研究者将目光集中到血浆和血清样本上来。血浆和血清样本的优点有多种[179]：首先，血浆和血清可能是最容易获得的生物样本，对于大动物来说，血浆和血清还是可以反复取样的生物样本。第二，血浆和血清不像其他生物样本需要纯化，或者要防止污染。第三，血浆和血清的研究是有相当长的历史的，远至希波克拉底时代的医学研究就将血液列为生命四要素之一。医学的历史有多长，血液的研究就有多长。第四，血浆和血清是非常好的蛋白质组研究目标，血浆和血清中不仅有传统意义上的“血液蛋白质”，而且有组织蛋白质（可以作为渗漏标志物研究）和数量惊人的多种多样的免疫球蛋白，这些蛋白质还有大量的修饰存在。因此，血浆和血清作为蛋白质组研究目标来说是非常合适的，蕴藏着巨大的信息资源。

血液样本可以分成全血、血浆和血清。全血中的可溶解部分就是血浆，血浆在发生凝结后，形成的上清部分是血清。当前的研究重点放在血浆和血清上，但是血清的形成过程中发生了体外的酶反应（凝血机制），其中涉及到了一系列的级联反应，对于生物样本来说不是令人放心的。而血浆要防止凝结，也要加入抗凝剂。所以，现在两方面的研究都很多，数据在一定程度上会有互补性，并且相互验证，成为当今体液蛋白质组学的研究主战场。

血浆和血清样本的研究也会碰到很大的困难[179]，血浆和血清中的蛋白质是非常丰富的，据估计，单单血清中就含有总计约 10 000 种的蛋白质。但丰度的差别是巨大的。

例如,每毫升血清中含有约 60～80mg 的蛋白质成分,而其中大半是白蛋白、免疫球蛋白、转铁蛋白、结合珠蛋白和脂蛋白等少数种类的蛋白质,种类上占绝大多数的低丰度蛋白质的含量与只占少数的极高丰度的蛋白质浓度差可达 $10^9 \sim 10^{10}$倍。这么高的比例如果不加处理,将严重影响对低丰度蛋白质的检测。另外,血浆和血清中占重要地位的糖基化蛋白质有为数众多的多态性。这些都是困扰学科发展而亟待解决的问题。

1. 诊断方向

血浆或血清蛋白质组学研究的一项目标是诊断方向,建立血浆和血清完整的蛋白质数据库,结合各种生理病理过程,发现更多的单一的或系列的生物学标志物。在此我们分样品前处理和样品分析两部分来讨论。

(1) 样品前处理

极高的浓度差和数以万种的蛋白质类型无疑是血清蛋白质组学的研究瓶颈,早期直接对血浆和血清样品进行双向凝胶电泳,获得 727 个蛋白质点,鉴定了其中 376 个,但是只是对应了 49 个蛋白质而已[180],其后的工作虽有所进步,但是在鉴定蛋白质的数量上没有大的提高[181]。人们开始一步步寻找解决办法,血浆和血清中有多种情况影响了对蛋白质组的分析,包括了白蛋白、转铁蛋白和免疫球蛋白的高含量,免疫球蛋白和糖基化蛋白质的多态性。直接进行分析是非常困难的,但是只要结合附加的分级措施,随后的分析就会相对容易,而且能够得到更好的结果。

蛋白质的分离技术已经发展了很多年,其中大多数都可以用来进行血浆和血清样品的分级处理。这些技术包括阴阳离子交换色谱、空间排阻色谱、亲和色谱、疏水色谱和反相色谱等,还包括平面自由流电泳和循环等电聚焦等方法。理论上,只要财力和物力可以接受,分级处理技术对血浆和血清样品分析的推进是没有限制的。研究者使用各种不同的分离策略,已在血浆和血清中鉴定了更多的蛋白质。

样品前处理的一种思路是:将样品尽可能多地分离成多个部分,使低丰度蛋白质在每一部分的相对含量得到增加,从而可能分析到低丰度的蛋白质;另一种思路是:将血浆和血清样品中的高含量的蛋白质尽可能除去,以消除它们在分析中带来的负面影响。但是后一种思路的实施有一定的难度,在血浆和血清样品中含量最高的蛋白质是白蛋白,它的含量高达 50%左右,但是从样品中分离白蛋白的努力很多都失败了。某些工作显示,使用超滤等方法试图除去高分子部分是徒劳的,这些努力要么是失败的,要么有丢失样品中其他重要低丰度蛋白质的可能[182～184]。但是在另一种高丰度蛋白质——免疫球蛋白的去除上倒是有成功的工作,使用的是交联着 A 蛋白和 G 蛋白的聚丙烯酰胺的亲和柱,在使用这种亲和层析前后的电泳可以看到,免疫球蛋白以及其重链和轻链浓度都大大降低了,而且将亲和柱上的蛋白质洗脱下来后的电泳显示柱上的蛋白质几乎全是免疫球蛋白以及其重链和轻链[185]。Joshua[185]等人将多维分离技术与质谱技术相结合,在方法学上进行了探索。具体方法:将取自健康女性的血清蛋白质样本先以蛋白质 A/G 的亲和色谱分离去除免疫球蛋白组分,胰蛋白酶酶解之后用强阳离子交换色谱将样品分成 120 个组分,每份组分分别进入反相 LC-MS/MS,最后使用 SEQUEST 软件分析数据。用这一方法,他们检测到共 490 种血清蛋白质,比未经分离步骤的血清样品中发现的蛋白质种类增加了 3～5 倍,而且还发现了一些丰度低至 ng/ml 的蛋白质,例如,人生长激素、白介素 12

和前列腺特异抗原。Joshua 等人还将 490 种血清蛋白质中的 134 种进行了归类，既在方法学上有所改进，又为血清蛋白质数据库的建立和可能的生物学标志物的发现做出了贡献。虽然这些工作已经取得了很大的突破，但是现在能够做的离真正的目标还很远，血浆和血清样品中的蛋白质总数远远超过现在所能检测的，而且其中还有很多的糖基化、磷酸化等多态性。总之，现在我们能检测的血浆和血清中的蛋白质和估计的蛋白质总数相比实在是相差悬殊。

本实验室也开展了使用循环等电聚焦和多维色谱技术进行样品分级的工作，取得了良好的结果，相应情况见有关章节。

另外，还有不经过任何前处理的大规模、超微量、高通量、全自动筛选蛋白质的技术路线，近来，美国 Ciphergen（赛弗吉）公司研制的 ProteinChip®Array 的仪器，并建立了一种新的蛋白质飞行质谱，即表面增强的激光解析离子化-飞行时间-质谱（surface-enhanced laser desorption/inionation-time of flight-mass spectra，SELDI-TOF-MS），这样的技术平台。SELDI-TOF-MS 已取得可喜的进展，筛选出了许多与疾病相关的新型生物学标志物，不仅为临床疾病的诊断和治疗等提供了新的选择，而且在基础研究、新药开发和疾病预防等方面具有广阔的应用前景。

(2) 样品分析

血浆和血清样品的分析可以大致的分成两种思路：一是前处理后进行双向凝胶电泳再进行质谱检测，二是直接进行混合物酶解再进行液质联用检测。

意大利学者在研究局灶性节段性肾小球硬化症（focal segmental glomerulo-sclerosis，FSGS，一种危险的青少年疾病，可导致严重的肾功能不全）时，使用的是前一种方法，采用先色谱后电泳的策略分析了病人的血浆样品，其色谱过程的基础是 A 蛋白、G 蛋白或 Br-CN 等交联的琼脂糖凝胶亲和色谱技术[186]。共纯化了 10 种差别表达的血浆蛋白质，质谱鉴定了其中 6 种，分别是白蛋白异构体、fibulin、载脂蛋白 J、玻基结合素（vitronectin）、纤维蛋白原 γ 链和植物凝集素相关的丝氨酸蛋白水解酶。

后一种方法是做所谓“shotgun”的质谱检测，就是混合物直接酶解后进行多维色谱分离并进行质谱检测。前面提到的 Joshua 等人的工作，样品分析时就采用了这种方法。他们共找到了 490 种蛋白质，这是在血清样品中找到蛋白质数量上的最高记录[185]。

不难想象，蛋白质组的图谱是很复杂的，这里包括经过或未经前处理的所有双向凝胶电泳的图谱和质谱的图谱，将其直接地、不加分析地使用在临床诊断中，是有相当大的困难的，同时这也超出了临床医生的能力范围。因此，临床蛋白质组的图谱识别就成为实际上的瓶颈问题。

蛋白质组的图谱识别是生物学中少有的、复杂的、带有类似数学性质的问题。但是，随着近年来各种相关技术和方法的发展，尤其是计算机技术的进步，原来一些难以在实验室中完成的计算过程也变得容易起来。另一方面，有越来越多的工作显示，数学中的某一些分支，如智能神经网络、丛分析和遗传算法等在临床蛋白质组工作中有很重要的使用价值。

应该承认，关于临床蛋白质组中的计算机和数学的内容是超出本书的范围的。但是，这些工作已经产生了巨大的成果，并可能对临床诊断学带来革命性的变化，从这些工作中看到的是蛋白质组学作为一门新学科和其他领域的科学结合的“化学反应”。美国的两大

医疗卫生机构 NCI 和 FDA 联合开展了对疾病诊断的研究，成立了临床蛋白质组计划(Clinical Proteomics Program)，并和计算机软件公司(Correlogic Systems Inc.)合作，在2002 年初在著名的医学杂志《柳叶刀》上发表了一篇重要的文章[187]。他们前所未有地将蛋白质组学和人工智能计算机技术整合到一起，应用到临床诊断中去，并取得了惊人的成果。根据文章的介绍，蛋白质组学技术产生的质谱图谱经过人工智能的软件自我学习模式的分析，诊断准确率非常高，阳性检出率达到 100%，对照组阴性检出率是 94%。试验中使用的是病人和健康人的血清，使用 SELDI-TOF 技术获得质谱图谱，开发的软件系统能够使用遗传算法和丛分析方法，对已知的癌症样本和对照健康人样本进行自学习过程，优化"丛类型"的识别，随着样本数量的增加，软件对癌症与健康两个"丛"的识别越来越准确，就可以运用到临床诊断中去了。可以认为，从基因组计划开始，数学和计算机技术已经深深地扎根于现代生物学当中，但是在蛋白质组领域内的这种突破性的结合才刚刚开始，这些技术将来会更多地应用在蛋白质组的其他方面。NCI 和 FDA 的科学家继续在其他癌症的诊断领域开展工作，不久之后，对前列腺癌的类似工作也取得了成功[188]。有着这些工作背后的巨大的商业价值的支撑，类似工作会迅速开展起来，难怪他们的负责人会说："We are extremely optimistic that this new approach will prove useful in detecting and diagnosing many other cancers and diseases in the future"。

2. 治疗方向

血浆或血清蛋白质组学研究的另一项目标是治疗方向，旨在筛选疾病相关抗原、发现自身抗体、找寻药物作用靶点和构建疫苗。当前较为流行的思路是血清学蛋白质组分析(serological proteome analysis，SERPA)路线，即研究整个致病微生物或疾病组织蛋白质组在血中所起的整体性体液免疫反应。这归功于几个方面的重大突破[189]：致病微生物(如疟原虫、弓浆虫、分枝杆菌、链球菌、沙门氏菌等)、疾病组织基因组学及蛋白质组学的成果，诸如蛋白质芯片这样的高通量、微量化、自动化、大规模的检测技术的建立，生物信息学对海量数据的整理、计算、分析、理解。我们可以用一幅基于 2DE 的实验设计(图23.4)[192]来概括 SERPA 的战略思想。

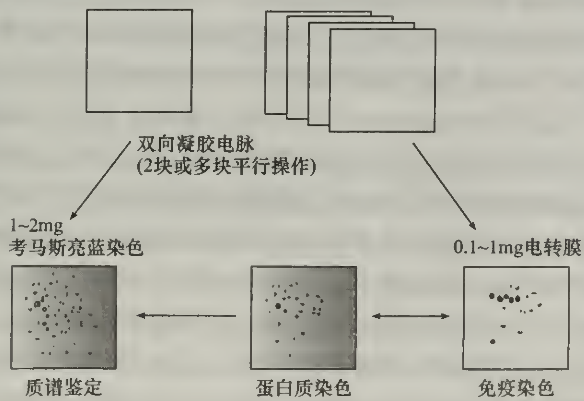


图 23.4　血清学蛋白质组分析路线

本着这样的思想，研究者们开展了一系列极具临床应用价值的工作。

Klade[190]等人先比较了健康肾脏与肾细胞癌肾脏组织的 2-DE 蛋白质图谱，再分别用病人及健康人的血清样本进行免疫杂交，结果发现肾细胞癌肾脏组织的 2-DE 蛋白质图谱上有 5 个特异性的点与病人血清样本发生免疫反应，提示这 5 个点的蛋白质可能是肿瘤相关抗原。用 Edman 降解法分析其中 2 个蛋白质的胰蛋白酶酶解肽段后，鉴别出这两种肿瘤相关抗原分别是平滑肌蛋白 22-α 和碳酸酐酶（carbonic anhydrase Ⅰ，CAⅠ）。Hanash[191]等人也以肺癌为模型做了相似的工作，旨在寻找与肿瘤相关抗原起反应的血清中的自身抗体。

Oresta[192]等人的工作比较系统，先用收集 74 个健康成人和 30 例感染了金黄色葡萄球菌病人的血清样本，用传统血清学方法分别筛选出 5 名高免疫血清样本金黄色葡萄球菌。再用培养的金黄色葡萄球菌（COL 菌株）制备蛋白质样品，做平行的 2-DE 实验，一部分上样量较低的 2-DE 胶与已选择的高免疫血清样本进行免疫杂交，ECL 方法选择出 21 个点进行胶内酶解结合 MS 鉴定；一部分上样量为毫克级的 2-DE 胶与 MS 相联系以鉴别目标蛋白质。MALDI-TOF-MS 共鉴定出 15 种已知或未知蛋白质作为细菌疫苗候选抗原。Lock[193]等人相同的办法，寻找到幽门螺杆菌中的 5 组候选抗原。

（二）有关脑脊液

脑脊液存在于脑室系统、脑周围的脑池和蛛网膜下腔内，被视为脑和脊髓的组织液和淋巴。脑脊液的主要功能是在脑、脊髓和颅腔、椎管之间起缓冲的作用，具有保护作用。脑浸浴于脑脊液中，由于浮力的作用，使脑的重量减轻到仅 50g 左右。另外，脑脊液还作为脑和血液之间进行物质交换的中介。成年人的脑脊液总量约 150ml。每天生成的脑脊液约 800ml，为脑脊液总量的 5～6 倍。但同时有等量的脑脊液被吸收入血液以较高的更新效率进行新陈代谢。正常脑脊液中含极微量蛋白质，且绝大部分是白蛋白，球蛋白仅有微量。在发生某些疾病时（如感染性脑膜炎脑脓肿、脑出血、脑瘤出血、蛛网膜下腔出血和外伤、脑血栓形成或栓塞、病毒性脑病、中毒性脑病等），脑脊液中的蛋白质可发生不同程度的量变和质变，且多为球蛋白。例如，波特兰-俄勒冈卫生科技大学（OHSU）的研究人员发现阿尔茨海默病与脑脊液里乳铁蛋白的含量升高是阿尔茨海默病恶化的一个征兆；用 ELISA 测定阿尔茨海默病、多发性梗死性痴呆及其他神经系统疾病患者脑脊液中 Tau 蛋白含量，证实脑脊液 Tau 蛋白的升高有助于阿尔茨海默病的诊断和鉴别诊断。

Townsend[194]（Oxford Glycosciences）以 2-DE 分离结合质谱识别技术，进行了旨在从脑脊液中发现阿尔茨海默病的蛋白质类疾病标志物的工作。通过对来源于 228 名病人总计 512 个血浆或脑脊液蛋白质样本的检验，发现关键性的血浆蛋白质（如白蛋白、转铁蛋白、免疫球蛋白）在血浆和脑脊液中的含量有显著的差异，似乎有选择机制将这些关键性的血浆蛋白质从脑脊液中移去却富集了结合珠蛋白等另一些种类的蛋白质；Townsend 所作的脑脊液蛋白表达谱共注释了 1 131 个点，潜在的阿尔茨海默病标志物就存在其中。

Davidsson[195]等人从突触病理学角度出发，运用蛋白质组学液相等电聚焦等手段研究不同病理时期病人脑脊液蛋白质图谱中的突触蛋白在神经精神病类疾病的作用。他们发现特殊种类的突触囊蛋白在神经胞外传导中发挥重要作用；突触前蛋白的作用涉及神经元生长、再生、萌芽和组织受损后的突触重组等多方面；患精神分裂症或阿尔茨海默病

时，许多种类的突触蛋白的表达情况都发生改变，这些疾病特异蛋白（disease-specific protein，DSP）为深入了解疾病病理、更好地进行早期诊断提供了线索；Rohlff[196]总结了迄今为止所发现的存在于各种急性或慢性中枢神经系统疾病的 30 余种 DSP，认为对某种神经系统失调疾病而言，特异而灵敏的标志物是一系列 DSP 而非某种单一蛋白质，这也正是蛋白质组学所反映的系统论思想的精髓所在。

应该指出，脑脊液是成分复杂的生物样本，尤其是其中富含多种盐类对相对少量蛋白质的分离有很大影响。Yuan[197]等人的工作很有特色，他们在走常规的 2-DE-MALDI-MS 路线之前，加入了三种不同的除盐方法制备脑脊液蛋白质样品：丙酮沉淀法、TCA/丙酮沉淀法和 Bio-Spin 柱预处理法。其中，Bio-Spin 柱是一种预制的脱盐柱，使用离心法，相对分子质量小于 6 000 的分子在离心时留在柱上，相对分子质量较大的蛋白质则可以通过。经对比，Bio-Spin 柱预处理法蛋白质损失量最低且发现蛋白质斑点数最多，是一种很值得推荐的样品脱盐方式。Yuan 等人选取病因不明的原发性后背痛（idiopathic low back pain）模式疾病，通过对病人腰穿进行脑脊液取样，经蛋白质组学研究，鉴定了 61 个点，反映的是 22 种蛋白质，其中 6 种蛋白质为首次报道，不仅丰富了脑脊液蛋白质数据库，而且提供了若干潜在的疾病标志物。

在辨别相近种类的神经精神疾病方面，Choe[198]等人做了探索性的工作。他们选取 Creutzfeld-Jakob disease（CJD）的两个类型：ante mortem variant Creutzfeld-Jakob disease（vCJD）和 ante mortem sporadic Creutzfeld-Jakob disease（spCJD）作为研究对象，分别取 2 例 vCJD 病人、3 例 spCJD 病人和 3 例对照用的脑脊液蛋白质样品，制备 2-DE 蛋白质表达图谱，运用丛分析的方案进行图谱识别，结果表明阿朴脂蛋白 E（apolipoprotein E）等 7 种脑脊液蛋白质是区别这两种 CJD 疾病的可能标志物。这一工作中脑脊液没有脱盐处理，因而 2-DE 的图谱不太理想，但能得到分析结果，看来丛分析在蛋白质组学中发挥的作用不容忽视。

脑脊液鼻漏是一种较常见的疾病，其发生主要是因为筛骨筛板和额窦后壁骨板甚薄，并与硬脑膜紧密相连，外伤时骨板与硬脑膜易同时破裂。脑脊液鼻漏可分为外伤性和非外性两种。外伤性以头部损伤为最常见病因。非外伤性主要由于颅内肿瘤、脑积水、颅底畸形或颅骨骨髓炎等引起。蛋白质组学的方法业已进行了脑脊液鼻漏的多分子标志物的检测，Burkhand[199]等用 18 例脑脊液鼻漏病人的鼻腔中脑脊液、鼻黏液和血浆的混合物制备蛋白质样本，进行 2-DE 等分析，找到了转铁蛋白、前列腺素 D 合成酶等已知或未知的 5 种蛋白质标志物。

另外，在脑脊液的蛋白质组学方法的探索道路上，一些研究小组也进行了有益的尝试。例如，Tseng[200]等人用填充了十二烷基磺酸钠（SDS）和多聚环氧乙烷（PEO）的毛细管进行电泳分离脑脊液或唾液的蛋白质样品取得较好的效果；先进的离子阱傅里叶转换离子回旋加速器共振质谱仪（electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry，ESI-FTICR-MS）也被用于包括脑脊液、血浆、唾液和尿液在内的多种体液的蛋白质组学研究[201]中。应用聚四氟乙烯（teflon）为 MALDI-TOF-MS 新的包埋基质的工作[202]也从脑脊液蛋白质的 2-DE 图谱中找到了 9 个潜在的标志物。美国 OGS 生物公司也不甘落后，利用先进的技术检测，分析了数百份阿尔茨海默病患者不同时期的脑脊液标本，尽管结果尚未正式公布，但已初步发现在该病发展过程中是由于蛋白

质表达异常。

总之，脑脊液的蛋白质组学研究在中枢神经系统疾病的蛋白质组学研究领域占有重要的一席，我们可以从图 23.5[196]看出脑脊液的蛋白质组学研究在这一领域的地位。

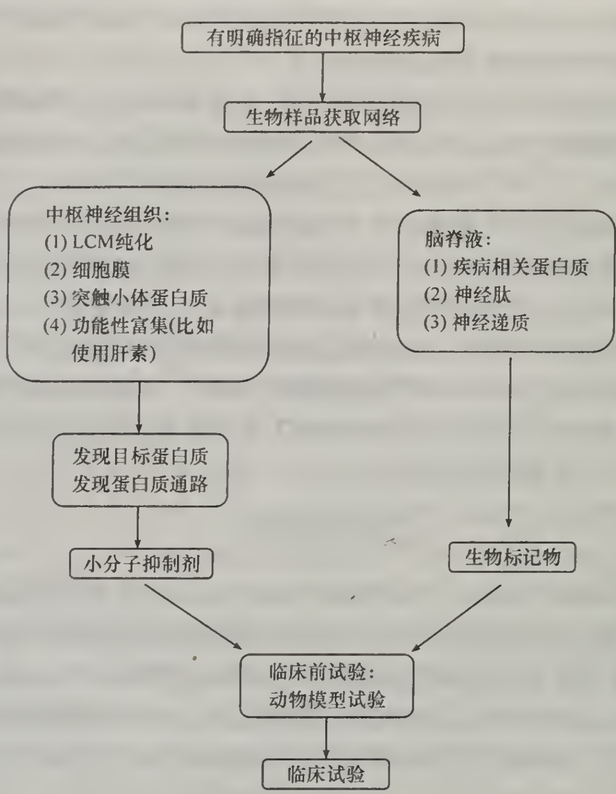


图 23.5 脑脊液的蛋白质组学与中枢神经系统疾病的蛋白质组学

（三）有关尿液

尿检是临床常见的检查，它是全身疾病检查的一部分，除检查常见的泌尿系统疾病（肾、肾盂、膀胱、尿道，炎症、肿瘤、结石）之外，还用来作为诊断糖尿病、肝病、血液病等疾病的参考指标。尿液的蛋白质组学研究的应用前景越发广阔。

最近，以肾脏范可尼综合征（renal Fanconi syndrome）为对象，运用在线的 nanoLC-MS/MS 技术，Cutillas[203]等人检测了来源于 3 例病人的 100 余份尿液的蛋白质样本，发现了若干（包括未知蛋白质在内）可能的生物活性多肽类标志物。考虑到尿液的成分复杂，Cutillas 等人设计了几步行之有效的样品处理过程：先用固相抽提（SPE）除盐，酶解后的肽段以强阳离子交换色谱柱（SCX）分离，SCX 所得的各组分再经反相液相色谱并进行在线的串联质谱分析，这样的实验设计体现了蛋白质组学中蛋白质、多肽及肽段需要充分分离的思想。

质谱免疫分析（mass sepctrometric immunoassay，MSIA）这一蛋白质组学技术综合了亲和捕获的选择性和质谱识别的灵敏性的优点，可进行完整蛋白质的辨认，是辨别蛋白质

突变体、翻译后修饰物及断裂产物等蛋白质变异体的利器。Kiernan[204]等人利用质谱免疫分析技术对比了尿液与血浆中的视黄醇结合蛋白质(retinol binding protein,RBP)图谱的异同,发现了一些全新的只在尿液中存在的视黄醇结合蛋白分解代谢产物。另外,他们还比较了健康人和肾脏损伤患者尿液 RBP 图谱的差异。

据报道,中风倾向的自发性高血压疾病(strok-prone spontaneously hypertensive disease,SHRSP)会引发高水平的蛋白尿。Sironi[205]等人以鼠为模式动物,运用 2-DE 追踪患病鼠尿液和血清中蛋白质表达情况,并与核磁共振成像(MRI)及组织学手段对比患病鼠脑部的变化,发现了先于 MRI 或组织学检测的因炎症反应、脉管渗透性改变所致的尿液蛋白质表达谱的改变(包括转铁蛋白、白蛋白、血液结合素、α_2-HS-糖蛋白、血管舒缓素结合蛋白等),这是从系统水平解释 SHRSP 引发蛋白尿的机制,同时也提供了一系列可能的疾病标志物。

SELDI-TOF 技术的商品化,使得面向蛋白质组学的尿检成为可能,Hampel[206]等人进行了初步的实验,他们对 20 只雄性 Sprague-Dawley 鼠和 25 名肾病患者进行追踪式的尿液取样,运用 SELDI-TOF 技术分析,分别在鼠及病人尿液中发现了 4 种不同相对分子质量的蛋白质的量变(鼠:9.9×10^3,18.7×10^3,21.0×10^3 和 66.3×10^3;人:9.75×10^3,11.75×10^3,23.5×10^3 和 66.4×10^3)。Vlahou[207]用 SELDI-TOF-MS 技术分析患者尿液中的蛋白质,发现多种蛋白质的差别表达,其中包括 5 种异常的新蛋白和 7 种异常的蛋白质组合,相对分子质量从 3.3×10^3 至 13.3×10^3 不等,从而把检测早期膀胱癌的准确率从泌尿细胞学法的 33%提高到 78%。这些结果再次证明 SELDI-TOF 技术是一种高灵敏、高通量、重复性好的痕量样品分析方法,应用前景广阔。

(四) 有关唾液

我国古代养生学十分重视唾液,用汉字表现,便是一个“活”字。古人认为,舌旁有水,即有唾液,才能成其“活”,可见唾液对于人的生命活动多么重要。现代医学认为,唾液中包含了血浆中的各种成分,包括黏蛋白、球蛋白等 10 多种酶、近 10 种维生素、多种矿物质以及有机酸和激素。比如,有一种叫唾液腺激素的物质能促进细胞的生长和分裂,并能加速核糖核酸蛋白质的合成;还有一种过氧化物酶,可以抑制致癌物质的毒性。因而,唾液的蛋白质组学研究也是极为必要的。

目前,唾液的蛋白质组学研究还处于起步阶段。Yao[208]等人以常规的 2-DE-MS 路线,研究了取自腮腺分泌物、颌下腺/舌下腺分泌物、唾液以及牙釉质(acquired enamel pellicle,俗称珐琅质,其蛋白质前体源于唾液)的蛋白质样品,发现唾液和牙釉质的蛋白质图谱比腮腺分泌物和颌下腺/舌下腺分泌物的蛋白质图谱更为复杂;虽然牙釉质的蛋白质前体源于唾液,但它们的蛋白质图谱有着显著的差异。这些结果对于数据库的建立以及深入认识口腔疾病大有裨益。

唾液蛋白质组学研究的另外一个研究方向与嗜血节肢动物有关,这类节肢动物的唾液内含有大量能严重影响宿主止血系统、炎症反应系统和免疫系统功能的物质,对这类节肢动物的唾液进行蛋白质组学研究也是揭示多种疾病发生发展的途径之一[209,210]。

(五) 其他

任何一种体液都反映相应细胞、组织或器官内外环境的情况,它们都可以成为蛋白质

组学的研究内容。除了上面介绍的几种体液的蛋白质组学研究工作外，乳汁乳头吸出液、胰液、胆汁、关节滑液等体液的蛋白质组学研究也开始了萌芽阶段的工作。

乳头吸出液(nipple aspirate fluid，NAF)蛋白质均为乳腺特异性的蛋白质，而且一般情况下比血液中相应蛋白质的浓度水平要高。Paweletz[211]等人以 SELDI-TOF 技术检测分析了 12 例乳腺癌患者和 15 例健康妇女的乳头吸出液，初步建立了乳头吸出液蛋白质表达数据库。随后，Sauter[212]等人也用同样的技术检测分析了 20 例乳腺癌患者和 13 例健康妇女的乳头吸出液，找出了 5 个差别表达的蛋白质，其中灵敏性及特异性良好(乳腺癌患者中检出率为 75%～84%，健康妇女仅为 0%～9%。有望成为乳腺癌标志物的是相对分子质量分别为 6 500 和 15 940 的两种蛋白质。

最近 Rosty[213]等人用含有铜离子的 IMAC-3 蛋白质生物芯片鉴定出一种相对分子质量约 16 570 的蛋白质，经鉴定，它是从急性胰腺炎和肝细胞癌患者的胰泡中释放出的一种名为 HIP/PAP-I(hepatocarcinoma-intestine-pancreas/ pancreatitis-associated-protein I)的蛋白质，它在胰腺癌患者的胰液中检出率为 67 %，而在其他类胰腺疾病中的检出率仅为 17%，具有较好的特异性。患者的胰液和血清中 HIP/PAP-I 水平与对照组相比的结果，都提示了可以依据患者胰液中 HIP/PAP-I 的含量来辅助鉴别胰腺癌。

由于胆汁成分复杂，Stark[214]等人精心设计了实验步骤(氯仿/甲醇抽提后，经甲醇/氯乙烯的反相色谱分离，最后氯仿/甲醇中进行凝胶过滤)，针对其中的疏水性蛋白质，进行 1D-SDS/PAGE-MS 分析，首次在胆汁中报道了血红蛋白 α 链、ATP 合成酶脂结合蛋白亚基 9、胰岛素样生长因子结合蛋白 7 的 N 端片段和单核细胞差异抗原 CD14 的中间片段。

目前，关节滑液(synovial fluid，SF)的常规临床检查主要是物理化学特征方面，包括透明度、黏度、葡萄糖含量、蛋白质浓度和细胞计数与形态观察，这些远不能满足对各类关节炎疾病的类型及病理时期精确检测的要求。关节滑液蛋白质组学研究正是区别不同疾病类型、发现疾病相关蛋白质的一个行之有效的研究手段。Andrea[215]等人选取 6 例风湿性关节炎、3 例骨关节炎和 1 例反应性关节炎病人的关节滑液样本，走 2-DE-MALDI-TOF-MS 的蛋白质组学常规路径，pH4-7 的 2-DE 图谱上显示约 300 个点，共鉴别了 150 个点，并对比了相应的病人血清蛋白质表达谱，发现纤维蛋白原 β 链降解产物只存在于关节滑液中，胞浆素 B 和 Calgranulin C 是风湿性关节炎病人关节滑液中特异表达的蛋白质，血清淀粉样 A 蛋白只在血清和风湿性关节炎病人关节滑液中显现。另外，此项工作还涉及了结合珠蛋白家族的疾病相关差异及个体间差异的比较分析。

三、展　　望

我们在此所列举的体液蛋白质组学的工作是有限的。一方面是因为各种体液类型不胜枚举，并且还有许多是蛋白质组学还未涉足的体液类型，如支气管肺泡液、围心腔液、内耳液、羊膜液、腹水和胸膜液，等等；另一方面，体液蛋白质组学与代谢组学(metablolomics)、肽组学(peptidomics)等相互交叉，又需要生物信息学(bioinformatics)、生物统计学(biostatistics)等多学科的支持，涉及内容之广已远远超出本节的范围。

循环系统将血液与全身各个组织器官密切联系，血液与相应部位的组织间液、细胞内

液直接或间接地进行着生生不息的体液交换。毫无疑问,在体液蛋白质组学研究中,重头戏是血浆蛋白质组学(包括血清蛋白质组学);当涉及特定器官、组织和细胞类型时,研究人员时常也会同时研究血浆(或血清)和对应的组织间液,以得到更为全面的信息。为适应体液蛋白质组学自身特点和发展需要,国际"人类蛋白质组织(HUPO)"于 2002 年 5 月中旬提出了 5 大研究开发计划,第一项计划就是以识别血液中较少量的蛋白质为目标的血浆蛋白质组计划(Plasma Proteome Project,PPP,见相关网址:www.hupo.org)。HUPO 为这一计划成立了由密歇根大学 Gilbert Omenn 教授领衔的专门的执行委员会。另外,下设样品收集与处理委员会、技术平台服务委员会、数据库与生物信息学委员会和样本银行等一系列分工细致、管理科学的功能机构。世界各地相关领域的几乎所有顶尖科学家和顶级研究小组都已投身到这项事业中来。我们有理由认为,大科学时代的体液蛋白质组学正在有组织、有计划、系统的国际化协作发展道路上稳步前进。

第五节　蛋白质组学在其他领域的应用

蛋白质组学是后基因组时代功能基因组研究的热点领域。蛋白质组学方法除了应用在微生物、肿瘤、神经系统、体液研究方面外,还广泛应用于植物、药物研究开发等方面。

一、蛋白质组学在植物研究中的应用

随着第一个植物基因组——拟南芥(*Arabidopis thaliana*)基因组测序工作的完成[216]和其他植物基因组和 EST 数据库的日益完善,植物蛋白质组的研究成为目前一个十分活跃的领域。

(一) 基于个体水平的植物蛋白质组学

1. 遗传多样性的研究

在研究株系、种群、栽培种或种之间的遗传多样性时,蛋白质组学方法已显示出巨大的优势。Barreneche 等人比较了欧洲 6 个国家的 23 株橡树的蛋白质组,结果表明 *Quercus petraea* 和 *Q. robur* 之间的遗传差异性不大[217]。David 等人比较了从单一种群繁衍而来但在不同环境下生长的小麦种群遗传多样性[218],表明所有的种群都与原先的种群不同,而绝大多数的差异性不是因为随机漂移,而是不同的气候条件导致。

2. 植物生理的研究

目前,该领域的蛋白质研究主要集中于胁迫条件下的生理变化,如干旱、高盐、ABA 处理等。胁迫会引起大量蛋白质在数量上和质量上的改变。Ramani 和 Apte 检测高盐胁迫下的水稻籽苗时,发现 35 种蛋白质被诱导表达,而 17 种蛋白质的表达受到抑制[219]。

3. 植物蛋白质组数据库

双向凝胶电泳中采用标准化、可重复的步骤,使得其适于大规模分析蛋白质。它可以与多种蛋白质鉴定方法(如蛋白质的免疫检测、质谱等)相结合,以获得关于蛋白质表达的

丰富信息。这些信息可在数据库中存储,表 23.6 列出了一些植物及植物组织的数据库。

(二) 基于组织和器官的植物蛋白质组学

植物基因的表达具有时空特异性,而不同组织和器官的蛋白质组学研究对于理解该特异性具有重大意义[220],这方面的相关报道较多。Tsuigita 等人[221]利用双向凝胶电泳技术分离水稻根、茎、叶、种子、芽、谷皮、愈合组织中的蛋白质,分析了 4 892 个点,通过 N 端测序鉴定了其中 3%。

Chang 等人[222]通过定量分析^{35}S-Met 的摄入量并结合 2-DE 和 MS 方法,比较低氧适应和缺氧条件下玉米根内的蛋白质合成。结果表明,在低氧适应的条件下,绝大多数蛋白质都能正常表达,只有少数几种蛋白质的表达选择性地增高。而在缺氧条件下,正常蛋白质合成受到抑制,并且没有出现其他蛋白质。共分析了 48 个蛋白质点,通过数据库搜索鉴定了其中 46 个。

Shen 等人通过双向凝胶电泳分离水稻叶鞘蛋白[223],考马斯亮蓝染色后检测到 352 个蛋白质点,确定了 31 种蛋白质的 N 端序列和 5 种蛋白质的中部序列。用基质辅助激光解吸飞行时间质谱(matrix-assisted laser desorption-ionization time-of-flight/mass spectrometry,MALDI-TOF/MS)分析了 55 种蛋白质,数据库搜索后发现,35 种蛋白质已知,而 20 种蛋白质是新发现的蛋白质。Komatsu 等人分析了悬浮培养的水稻细胞的蛋白质[224],双向凝胶电泳分离得到 103 个蛋白质点,确定了 20 种蛋白质的 N 端序列和 32 种蛋白质的中部序列。确定序列的蛋白质与蛋白质数据库中已知蛋白质比对后,发现一些蛋白质具有重要的生理功能,如钙结合蛋白、超氧化物歧化酶、抗坏血酸过氧化物酶等。相关数据库见表 23.6。

表 23.6　植物相关数据库

植物种类或组织	网　　址
海岸松	http://www.pierroton.inra.fr/genetics/2D/
马铃薯	http://proteomics.cancer.dk/cgi-bin/gel-launcher? Potato+(IEF)£potato 玉米
	http://aestivum.moulon.inra.fr/cgi-bin/imgd/dbprot/dbprot.html
小麦	http://www-lecb.ncifcrf.gov/2dwgDB/#searching
水稻花粉囊	http://semele.anu.edu.au/2d/2d.html
橡树根组织	http://www.pierroton.inra.fr/genetics/2D/oak-root.jpg
桉树木质部	http://www.pierroton.inra.fr/genetics/2D/xylemeeuca.jpg

(三) 亚细胞水平的植物蛋白质组学

已知的亚细胞水平的植物蛋白质组学数据库包括植物细胞膜蛋白质组数据库(http://spinx.rug.ac.be:8080/ppmdb/index.html)、拟南芥线粒体蛋白质组数据库(http://www.gartenbau.uni-hannover.de/genetik/AMPP)等。目前在植物中研究得最多的是叶绿体。据估计,在高等植物中,有21 000～25 000 种蛋白质[225],而叶绿体蛋白质占 10%～25%[226],可见其在植物细胞中的重要性。

Peltier 等人运用蛋白质组学方法分析了豌豆叶绿体的内囊体的可溶和外周蛋白质[227]。他们估计至少存在 200～230 种内腔和外周蛋白,不包括同种型和可能的翻译后修饰,鉴定了其中 61 种蛋白质。这些蛋白质中的 33 种蛋白质已知功能或功能结构域。Yamagushi 和 Subramanian 鉴定了菠菜叶绿体中核糖体 30S 和 50S 亚基的蛋白质[228]。

(四) 植物蛋白质组学研究的局限性和可能性

Edman 测序或质谱方法对蛋白质的鉴定,较大程度上依赖于相关植物基因组和 EST 数据库的充实和发展。因此目前存在的问题是:对于未进行基因组测序的植物来说,植物蛋白质组学研究在何种程度上是可行的。有两个选择:①如果 EST 序列已知,那么可通过由质谱方法获得的数据来鉴定蛋白质,但有时许多蛋白质未被 EST 数据库所包含;②同源比对的方法,相当数量的植物蛋白质之间存在着同一/同源性,获得相应蛋白质的氨基酸序列信息,利用 FASTA 或相似的搜索引擎进行数据库搜索。MALDI-TOF 在产生数量巨大的肽段时,可以提高同源比对的成功率。如 MALDI-TOF 应用于豌豆叶绿体蛋白质的研究[227]。

二、蛋白质组学在药物开发方面的研究

人类基因组序列的完成,加快了药物研究的进程,但仅仅通过测定基因序列并不能确定药物作用的靶点。因为蛋白质是生命活动的执行者,基因序列与疾病的关系是间接的;另一方面,蛋白质蕴含的信息远比核酸丰富。因此,蛋白质组学是基因组学和药物发现之间的桥梁,它不仅可能提供治疗的靶点,如相互作用蛋白质组学提供了靶分子与其他分子的关系;也可能提高药物发现下游过程的效率,如结构蛋白质组提供的蛋白质空间构象将促进药物分子的设计。

药物蛋白质组学研究的内容包括:发现所有可能的药物作用靶点,以及针对这些靶点的全部可能的化合物;应用蛋白质组学方法研究药物作用机制和毒理学;发现药物作用的特异蛋白质作为患者选择有效药物的依据和临床诊断的标志物[229]。

一方面,利用药物蛋白质组学方法可以进行靶点的探测。通过正常个体和患病个体的蛋白质组学分析,可以鉴定出在患病状态时表达异常的蛋白质。如组织蛋白酶 K 仅在破骨细胞中特异性表达,因此抑制组织蛋白酶 K 可能对骨质疏松症有治疗作用[230]。许多疾病与药物传导途径异常有关,通过对受调节的基因及其产物的分析研究,可以推测其受体介导的生化代谢的改变,因此信号分子可以考虑作为治疗药物设计的靶点[231]。如 Wnt 信号通路中结肠癌抑制因子(APC)的突变将引起结肠癌[232]。

另一方面,蛋白质组学还可用于药物毒性的研究。正常组织或已知毒性试剂处理的组织蛋白质表达情况的比较,可以表明药物的毒性及活性。早在 20 世纪 80 年代 Anderson 等就开始利用双向凝胶电泳进行毒理学的研究[233,234],并在多篇文章中提到,“此技术可产生鉴定组织损伤的生物标志,提供检测毒物的方法。”蛋白质组学在毒理学领域成功应用的例子就是对环孢菌素 A(cyclosporine,CsA)效应的研究[235,236]。CsA 是用于器官移植的临床免疫抑制剂,但具有肾脏毒性的副作用。Steiner 等人通过 CsA 对小鼠肾脏组织影响的蛋白质组学研究[235],首次阐明了 CsA 肾毒性机制。在用药后的小鼠肾脏组织

的蛋白质组中，钙结合蛋白 calbindin D_{28}（细胞内的一种钙转运蛋白）明显减少，这种减少伴随着尿钙排泄物和皮层脊髓管内石灰化的增加，这解释了用该药后肾小管的钙累积和由此造成的肾小管毒性。

三、结　　语

蛋白质组学是在蛋白质水平上大规模研究基因功能的强有力的工具。随着蛋白质组学技术的进一步发展以及相关数据库的进一步充实，蛋白质组学会有更广阔的发展前景，将对现代生物学及生物技术的发展做出重大贡献。

（丁士健　唐柳娅　李　辰）

参考文献

1 www.shenliu.com/weishengwu/gongxing.htm

2 http://www.37c.com.cn/literature/library/theory/007/00701001.html

3 http://www.ebi.ac.uk/

4 Maureen M. Barr, super models. Physiol Genom, 2003, 13:15～24

5 Doyle SA, Murphy MB, Massi JM *et al*. High-throughput proteomics: a flexible and efficient pipeline for protein production. J Proteome Res, 2002, 1(6):531～536

6 Rosen R, Ron EZ. Proteome analysis in the study of the bacterial heat-shock response. Mass Spectrom Rev, 2002, 21(4):244～265

7 Huang YJ, Swapna GV, Rajan PK *et al*. Solution NMR structure of ribosome-binding factor A (RbfA), a cold-shock adaptation protein from Escherichia coli. J Mol Biol, 2003, 327 (2):521～536

8 Savchenko A, Yee A, Khachatryan A *et al*. Strategies for structural proteomics of prokaryotes: quantifying the advantages of studying orthologous proteins and of using both NMR and X-ray crystallography approaches. Proteins, 2003, 50(3):392～399

9 Gevaert K, Van Damme J, Goethals M *et al*. Chromatographic isolation of methionine-containing peptides for gel-free proteome analysis: identification of more than 800 Escherichia coli proteins. Mol Cell Proteomics, 2002, 1(11):896～903

10 Cruickshank WH, Malchy BL, Kaplan H. Diagonal chromatography for the selective purification of tyrosyl peptides. Can J Bio-chem, 1974, 52:1013～1017

11 Mirgorodskaya OA, Kozmin YP, Titov MI *et al*. Quantitation of peptides and proteins by matrix-assisted laser desorption/ionization mass spectrometry using ^{18}O-labeled internal standards. Rapid Commun Mass Spectrom, 2000, 14:1226～1232

12 Yan JX, Devenish AT, Wait R *et al*. Fluorescence two-dimensional difference gel electrophoresis and mass spectrometry based proteomic analysis of Escherichia coli. Proteomics, 2002, 2(12):1682～1698

13 Wojcik J, Boneca IG, Legrain P. Prediction, assessment and validation of protein interaction maps in bacteria. J Mol Biol, 2002, 323(4):763～770

14 Wojcik J, Schä chter V. Protein-protein interaction map inference using interacting domain profile pairs. Bioinformatics, 2001, 17:296～305

15 Ge Y, Lawhorn BG, ElNaggar M *et al*. Top down characterization of larger proteins (45 kDa) by electron capture dissociation mass spectrometry. J Am Chem Soc, 2002, 124(4):672～678

16 Pazos F, Valencia A. Similarity of phylogenetic trees as indicator of protein-protein interaction. Protein Eng, 2001, 14(9):609～614

17 Wang S, Regnier FE. Proteomics based on selecting and quantifying cysteine containing peptides by covalent chromatography. J Chromatogr A, 2001, 924(1-2): 345～357
18 Wang S, Regnier FE. Proteolysis of whole cell extracts with immobilized enzyme columns as part of multidimensional chromatography. J Chromatogr A, 2001, 913(1-2): 429～436
19 Jelinsky SA, Estep P, Church GM *et al*. Regulatory networks revealed by transcriptional profiling of damaged Saccharomyces cerevisiae cells: Rpn4 links base excision repair with proteasomes. Mol Cell Biol, 2000, 20: 8157～8167
20 Wyrick JJ, Aparicio JG, Chen T *et al*. Genome-wide distribution of ORC and MCM proteins in S. cerevisiae: high-resolution mapping of replication origins. Science, 2001, 294: 2357～2360
21 Uetz P. Two-hybrid arrays. Curr Opin Chem Biol, 2002, 6: 57～62
22 Kumar A, Snyder M. Protein complexes take the bait. Nature, 2002, 415: 123～124
23 Kumar A, Agarwal S, Heyman JA *et al*. Subcellular localization of the yeast proteome. Genes Dev, 2002, 16: 707～719
24 Zhu H, Bilgin M, Bangham R *et al*. Global analysis of protein activities using proteome chips. Science, 2001, 293: 2101～2105
25 Perego P, Jimenez GS, Gatti L *et al*. Yeast mutants as a model system for identification of determinants of chemosensitivity. Pharmacol Rev, 2000, 52(4): 477～492
26 Feldhaus MJ, Siegel RW, Opresko LK *et al*. Flow-cytometric isolation of human antibodies from a nonimmune Saccharomyces cerevisiae surface display library. Nat Biotechnol, 2003, 21(2): 163～170
27 ftp://ftp.mshri.on.ca/pub/BIND/Tools/MCODE
28 Bader GD, Hogue CW. An automated method for finding molecular complexes in large protein interaction networks. BMC Bioinformatics, 2003, 4(1): 2
29 Krogan NJ, Kim M, Ahn SH *et al*. RNA polymerase II elongation factors of Saccharomyces cerevisiae: a targeted proteomics approach. Mol Cell Biol, 2002, 22(20): 6979～6992
30 Yoon HJ, Feoktistova A, Wolfe BA *et al*. Proteomics analysis identifies new components of the fission and budding yeast anaphase-promoting complexes. Curr Biol, 2002, 12(23): 2048～2054
31 Ohi MD, Link AJ, Ren L *et al*. Proteomics analysis reveals stable multiprotein complexes in both fission and budding yeasts containing Myb-related Cdc5p/Cef1p, novel pre-mRNA splicing factors, and snRNAs. Mol Cell Biol, 2002, 22(7): 2011～2024
32 Ho Y, Gruhler A, Heilbut A *et al*. Systematic identification of protein complexes in Saccharomyces cerevisiae by mass spectrometry. Nature, 2002, 415(6868): 180～183
33 Gavin A C, Bösche M *et al*. Functional organization of the yeast proteome by systematic analysis of protein complexes. Nature, 2002, 415(6868): 141～147
34 Tong A H Y, Drees B, Nardelli G *et al*. A combined experimental and computational strategy to define protein interaction networks for peptide recognition modules. Science, 2002, 295: 321～324
35 Smolka M, Zhou H, Aebersold R. Quantitative protein profiling using two-dimensional gel electrophoresis, isotope-coded affinity tag labeling, and mass spectrometry. Mol Cell Proteomics, 2002, 1(1): 19～29
36 Smith RD, Anderson GA, Lipton MS *et al*. The use of accurate mass tags for high-throughput microbial proteomics. OMICS, 2002, 6(1): 61～90
37 Loughrey Chen S, Huddleston MJ, Shou W *et al*. Mass spectrometry-based methods for phosphorylation site mapping of hyperphosphorylated proteins applied to Net1, a regulator of exit from mitosis in yeast. Mol Cell Proteomics, 2002, 1(3): 186～196
38 Bertone P, Kluger Y, Lan N *et al*. SPINE: an integrated tracking database and data mining approach for identifying feasible targets in high-throughput structural proteomics. Nucleic Acids Res, 2001, 29 (13): 2884～2898
39 http://www.zgxl.net
40 Ren SX, Fu G, Jiang XG *et al*. Unique physiological and pathogenic features of Leptospira interrogans revealed by whole-genome sequencing. Nature, 2003, 422: 888～893
41 Jungblut PR, Grabher G, Stoffler G. Comprehensive detection of immunorelevant Borrelia garinii antigens by two-dimen-

sional electrophoresis. Electrophoresis,1999,20(18):3611~3622

42 Demirev PA, Lin JS, Pineda FJ *et al*. Bioinformatics and mass spectrometry for microorganism identification: proteome-wide post-translational modifications and database search algorithms for characterization of intact H. pylori. Anal Chem, 2001,73(19):4566~4573

43 Reid SD, Green NM, Sylva GL *et al*. Postgenomic analysis of four novel antigens of group a streptococcus: growth phase-dependent gene transcription and human serologic response. J Bacteriol,2002,184: 6316~6324

44 Florens L, Washburn MP, Raine JD *et al*. A proteomic view of the Plasmodium falciparum life cycle. Nature,2002, 419:520~526

45 Lasonder E, Ishihama Y, Andersen JS *et al*. Analysis of the Plasmodium falciparum proteome by high-accuracy mass spectrometry. Nature,2002,419: 537~542

46 Govorun VM, Moshkovskii SA, Tikhonova OV *et al*. Comparative analysis of proteome maps of Helicobacter pylori clinical isolates. Blectrophoresis,2003,68(1):42~49

47 Chu PW, Yap MN, Wu CY *et al*. A proteomic analysis of secreted proteins from xylan-induced Bacillus sp strain K-1 . Electrophoresis,2000,21(9):1740~1745

48 Young NM, Brisson JR, Kelly J *et al*. Structure of the N-linked glycan present on multiple glycoproteins in the Gram-negative bacterium, Campylobacter jejuni. J Biol Chem,2002,277 (45):42530~42539

49 Kahng HY, Cho K, Song SY *et al*. Enhanced detection and characterization of protocatechuate 3,4-dioxygenase in Acinetobacter lwoffii K24 by proteomics using a column separation. Biochem Biophys Res Commun,2002,295(4): 903~909

50 Haney SA, Alksne LE, Dunman PM *et al*. Genomics in anti-infective drug discovery—getting to endgame. Curr Pharm Des,2002,8(13):1099~1118

51 Lock RA, Cordwell SJ, Coombs GW *et al*. Proteome analysis of Helicobacter pylori: major proteins of type strain NCTC 11637,Pathology,2001,33(3):365~374

52 Antelmann H, Yamamoto H, Sekiguchi J *et al*. Stabilization of cell wall proteins in Bacillus subtilis: a proteomic approach. Proteomics,2002,2:591~602

53 Haas G, Karaali G, Ebermayer K *et al*. Immunoproteomics of Helicobacter pylori infection and relation to gastric disease. Proteomics,2002,2:313~324

54 Vytvytska O, Nagy E, Bluggel M *et al*. Identification of vaccine candidate antigens of Staphylococcus aureus by serological proteome analysis. Proteomics,2002,2:580~590

55 Chakravarti DN, Fiske MJ, Fletcher LD *et al*. Application of genomics and proteomics for identification of bacterial gene products as potential vaccine candidates. Vaccine,2000,19(6):601~612

56 Klade CS. Proteomics approaches towards antigen discovery and vaccine development. Curr Opin Mol Ther,2002,4: 216~223

57 Hendrickson RC, Douglass JF, Reynolds LD *et al*. Mass spectrometric identification of mtb81,a novel serological marker for tuberculosis. J Clin Microbiol,2000,38: 2354~2361

58 Klade CS, Voss T, Krystek E *et al*. Identification of tumor antigens in renal cell carcinoma by serological proteome analysis. Proteomics,2001,1:890~898

59 Poon TC, Johnson PJ. Proteome analysis and its impact on the discovery of serological tumor markers. Clin Chim Acta, 2001,313: 231~239

60 Pederson T. The immunome. Mol Immunol,1999,36: 1127~1128

61 Tenover FC, Biddle JW, Lancaster MV. Increasing resistance to vancomycin and other glycopeptides in Staphylococcus aureus. Emerg Infect Dis,2001,7:327~332

62 Shaw AC, Larsen MR, Roepstorff P *et al*. Identification and characterization of a novel Chlamydia trachomatis reticulate body protein. FEMS Microbiol Lett,2002,212(2):193~202

63 Molestina RE, Klein JB, Miller RD *et al*. Proteomic analysis of differentially expressed Chlamydia pneumoniae genes during persistent infection of HEp-2 cells. Infect Immun,2002,70(6):2976~2981

64 Regula JT, Ueberle B, Boguth G *et al*. Towards a two-dimensional proteome map of Mycoplasma pneumoniae. Electrophoresis, 2000, 21(17):3765~3780

65 Ueberle B, Frank R, Herrmann R. The proteome of the bacterium Mycoplasma pneumoniae: comparing predicted open reading frames to identified gene products. Proteomics, 2002, 2 (6):754~764

66 Wasinger VC, Pollack JD, Humphery-Smith I. The proteome of Mycoplasma genitalium. Chaps-soluble component. Eur J Biochem, 2000, 267(6):1571~1582

67 Balasubramanian S, Schneider T, Gerstein M *et al*. Proteomics of Mycoplasma genitalium: identification and characterization of unannotated and atypical proteins in a small model genome. Nucleic Acids Res, 2000, 28(16):3075~3082

68 Bruneau JM, Maillet I, Tagat E *et al*. Drug induced proteome changes in Candida albicans: comparison of the effect of beta(1,3) glucan synthase inhibitors and two triazoles, fluconazole and itraconazole. Proteomics, 2003, 3(3):325~336

69 Graves PR, Kwiek JJ, Fadden P *et al*. Discovery of novel targets of quinoline drugs in the human purine binding proteome. Mol Pharmacol, 2002, 62(6):1364~1372

70 Cohen AM, Rumpel K, Coombs GH *et al*. Characterisation of global protein expression by two-dimensional electrophoresis and mass spectrometry: proteomics of Toxoplasma gondii. Int J Parasitol, 2002, 32(1):39~51

71 Rout MP, Field MC. Isolation and characterization of subnuclear compartments from Trypanosoma brucei-Identification of a major repetitive nuclear lamina component. Journal of Biol Chem, 2001, 276(41):38261~38271

72 Pearson TW. Procyclins, proteases and proteomics: dissecting trypanosomes in the tsetse fly. Trends Microbiol, 2001, 9 (7):299~301

73 Sam Hanash. Disease proteomics. Nature, 2003, 422(13):226~232

74 Tang ZY, Yu YQ, Zhou XD. An important approach to prolonging survival futher after radical resection of AFP positive hepatocellular carcinoma. J Exp Clin Cancer Res, 1984, 3:359~366

75 Lander ES, Linton LM, Birren B *et al*. Initial sequencing and analysis of the human genome. Nature, 2001, 409:860~921

76 Velculescu VE, Zhang L, Vogelstein B *et al*. Serial analysis of gene expression. Science, 1995, 270:484~487

77 Schena M, Shalon D, Davis RW *et al*. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. Science, 1995, 270: 467~470

78 Iyer VR, Eisen MB, Ross DT *et al*. The transcriptional program in the response of human fibroblasts to serum. Science, 1999, 283:83~87

79 Anderson NL, Anderson NG. Proteome and proteomics: new technologies, new concepts, and new words. Electrophoresis, 1998, 19:1853~1861

80 Abbott A. Proteomics: The society of proteins. Nature, 2002, 417:894~896

81 Stoeckli M, Chaurand P, Hallahan DE *et al*. Imaging mass spectrometry: a newtechnology for the analysis of protein expression in mammalian tissues. Nat Med, 2001, 7: 493~496

82 Page MJ, Amess B, Townsend RR *et al*. Proteomic definition of normal luminal and myoepithelial breast cells purified from reduction mammoplasties. Proc Natl Acad Sci USA, 1999, 96:12589~12594

83 Reid GE, Rasmussen RK, Dorow DS *et al*. Capillary column chromatography improves sample preparation for mass spectrometric analysis: complete characterization of human alpha-enolase from two-dimensional gels following in situ proteolytic digestion. Electrophoresis, 1998, 19:946~955

84 Rasmussen RK, Ji H, Eddes JS *et al*. Two-dimensional electrophoretic analysis of human breast carcinoma proteins: mapping of proteins that bind to the SH3 domain of mixed linkage kinase MLK2, Electrophoresis, 1997, 18:588~598

85 Chong BE, Yan F, Lubman DM *et al*. Chromatofocusing nonporous reversed-phase high-performance liquid chromatography/ electrospray ionization time-of-flight mass spectral analysis. Rapid Commun. Mass Spectrom, 1999, 13:1808~1812

86 Bergman AC, Benjamin T, Alaiya AA *et al*. Identification of gel-separated tumor marker proteins by mass spectrometry. Electrophoresis, 2000, 21:679~686

87 Gygi SP, Khan DKM, Gingras AC *et al*. Protein analysis by mass spectrometry and sequence database searching: tools

for cancer research in the post-genomic era. Electrophoresis, 1999, 20: 310～318

88 Vercoutter-Edouart AS, Lemoine J, Le Bourhis X *et al*. Proteomic analysis reveals that 14-3-3 sigma is down-regulated in human breast cancer cells. Cancer Res, 2001, 61: 76～80

89 Rao S, Aberg F, Nieves E *et al*. Identification by mass spectrometry of a new α-tubulin isotype expressed in human breast and lung carcinoma cell lines. Biochemistry, 2001, 40: 2096～2103

90 Watkins B, Szaro R, Ball S *et al*. Detection of early-stage cancer by serum protein analysis. Am Lab, 2001, 31: 32～36

91 Petricoin EF, Ardekani AM, Hitt BA *et al*. Use of proteomic patterns in serum to identify ovarian cancer. Lancet, 2002, 359: 572～577

92 Sullivan CM, Smith DM, Matsui NM *et al*. Identification of constitutive and gamma-interferon-and interleukin 4-regulated proteins in the human renal carcinoma cell line ACHN. Cancer Res, 1997, 57: 1137～1143

93 Ostergaard M, Wolf H, Orntoft TF *et al*. Psoriasin (S100A7): a putative urinary marker for the follow-up of patients with bladder squamous cell carcinomas. Electrophoresis, 1999, 20: 349～354

94 Celis JE, Ostergaard M, Rasmussen HH *et al*. A comprehensive protein resource for the study of bladder cancer. Electrophoresis, 1999, 20: 300～309

95 Celis JE, Wolf H, Ostergaard M. Bladder squamous cells carcinoma biomarkers derived from proteomics. Electrophoresis, 2000, 21: 2115～2121

96 Vlahou A, Schellhammer PF, Mendrinos S *et al*. Development of a novel proteomic approach for the detection of transitional cell carcinoma of the bladder in urine. Am J Pathology, 2001, 158: 1491～1502

97 Xiao Z, Jiang XM, Beckett ML *et al*. Generation of a baculovirus recombinant prostate-specific membrane antigen and its use in the development of a novel protein biochip quantitative immunoassay. Protein Exp Purif, 2000, 19: 12～21

98 Wang S, Diamond DL, Hass GM *et al*. Identification of prostate specific membrane antigen (PSMA) as the target of monoclonal antibody 107-1A4 by proteinchip array, surface-enhanced laser desorption/ionization (SELDI) technology. Int J Cancer, 2001, 92: 871～876

99 Nelson PS, Han D, Rochon Y *et al*. Comprehensive analyses of prostate gene expression: convergence of expressed sequence tag databases, transcript profiling and proteomics. Electrophoresis, 2000, 21: 1823～1831

100 Emmert-Buck MR, Gillespie JW, Paweletz CP *et al*. An approach to proteomic analysis of human tumors, Mol Carcinog, 2000, 27: 158～165

101 Paweletz CP, Gillespie JW; Ornstein DK *et al*. Rapid protein display profiling of cancer progression directly from human tissue using a protein biochip. Drug Develop Res, 2000, 49: 34～42

102 Lawrie LC, Curran S, McLeod HL *et al*. Application of laser capture microdissection and proteomics in colon cancer, Mol Pathol, 2001, 54: 253～258

103 Simpson RJ, Connolly LM, Eddes JS *et al*. Proteomic analysis of the human colon carcinoma cell line (LIM 1215): development of a membrane protein database. Electrophoresis, 2000, 21: 1707～1732

104 Stulik J, Osterreicher J, Koupilova K *et al*. The analysis of S100A9 and S100A8 expression in matched sets of macroscopically normal colon mucosa and colorectal carcinoma: the S100A9 and S100A8 positive cells underlie and invade tumor mass. Electrophoresis, 1999, 20: 1047～1054

105 Stulik J, Koupilova K, Osterreicher J *et al*. Protein abundance alterations in matched sets of macroscopically normal colon mucosa and colorectal carcinoma. Electrophoresis, 1999, 20: 3638～3646

106 Cole AR, Ji H, and Simpson RJ. Proteomic analysis of colonic crypts from normal, multiple intestinal neoplasia and p53-null mice: a comparison with colonic polyps. Electrophoresis, 2000, 21: 1772～1781

107 Melis R, White R. Characterization of colonic polyps by two-dimensional gel electrophoresis. Electrophoresis, 1999, 20: 1055～1064

108 Sinha P, Hutter G, Kottgen E *et al*. Search for novel proteins involved in the development of chemoresistance in colorectal cancer and fibrosarcoma cells in vitro using two-dimensional electrophoresis, mass spectrometry and microsequencing. Electrophoresis, 1999, 20: 2961～2969

109 Liotta L, Petricoin EF. Molecular profiling of human cancer. Nat Rev Genet, 2000, 1: 48～56

110 Seow TK, Ong SE, Liang RCM *et al*. Two-dimensional electrophoresis map of the human heptatocellular carcinoma cell line, HCC-M, and identification of the separated proteins by mass spectrometry. Electrophoresis, 2000, 21:1787~1813
111 Yu LR, Zeng R, Shao XX *et al*. Identification of differentially expressed proteins between human hepatoma and normal liver cell line s by two-dimensional electrophoresis and liquid chromatography-ion trap mass spectrometry. Electrophoresis, 2000, 21:3058~3068
112 Sinha P, Hutter G, Kottgen E *et al*. Increased expression of epidermal fatty acid binding protein, cofilin, and 14-3-3-σ (stratifin) detected by two-dimensional gel electrophoresis, mass spectrometry and microsequencing of drug-resistant human adenocarcinoma of the pancreas. Electrophoresis, 1999, 20:2952~2960
113 Chuman Y, Bergman AC, Ueno T *et al*. A member of the aspertic protease family, is abundantly expressed in normal lung and kidney tissue and is expressed in lung adenocarcinomas. FEBS Lett, 2000, 462:129~134
114 Gamble SC, Dunn MJ, Wheeler CH *et al*. Expression of proteins coincident with inducible radioprotection in human lung epithelial cells. Cancer Res, 2000, 60:2146~2151
115 Von Eggeling F, Davies H, Lomas L *et al*. Tissue-specific microdissection coupled with protein chip array technologies: applications in cancer research. Biotechniques, 2000, 29: 1066~1069
116 Ferrari L, Seraglia R, Rossi CR *et al*. Protein profiles in sera of patients with malignant cutaneous melanoma. Rapid Commum Mass Spectrom, 2000, 14:1149~1154
117 Sinha P, Kohl S, Fischer J *et al*. Identification of novel proteins associated with the development of chemoresistance in malignant melanoma using two-dimensional electrophoresis. Electrophoresis, 2000, 21:3048~3057
118 Bjellerup P, Theodorsson E, Jornvall H *et al*. Limited neuropeptide Y precursor processing in unfavourable metastic neuroblastoma tumors. Br J Cancer, 2000, 83:171~176
119 Yu LR, Shao XX, Jiang WL *et al*. Proteome alterations in human hepatoma cells transfected with antisense epidermal growth factor receptor sequence. Electrophoresis, 2001, 22:3001~3008
120 Pepe MS, Etzioni R, Feng Z *et al*. Phases of biomarker development for early detection ofcancer. J Natl Cancer Inst, 2001, 93:1054~1056
121 Wulfkuhle JD, Liotta LA, Petricoin EF. Proteomic applications for the early detection of cancer. Nat Rev Cancer, 2003, 3:267~275
122 Hyden H, Lange PW. Protein synthesis during learning. Science, 1969, 164:200~201
123 Hyden H, Lange PW. Protein synthesis in the hippocampal pyramidal cells of rats during a behavioral test. Science, 1968, 159:1370~1373
124 Cotman CW, Mathews DA. Synaptic plasma membranes from rat brain synaptosomes: isolation and partial characterization. Biochim Biophys, 1971, 249:380~394
125 Gombos G, Morgan I, Breckenridge WC *et al*. Isolation and biochemical structure of synaptosomal membrane. C R Seances Soc Biol Fil, 1971, 165: 499~506
126 Waehneldt TV, Morgan IG, Gombos G. The synaptosomal plasma membrane: protein and glycoprotein composition. Brain Res, 1971, 34:403~406
127 Burden S. Identification of an intracellular postsynaptic antigen an the frog neuromuscular junction. J Cell Biol, 1982, 94:521~530
128 Gurd JW, Gordon-Weeks PR, Evans WH. Identification and localization of concanavalin A binding sites on isolated postsynaptic densities. Brain Res, 1983, 276:141~146
129 Froehner SC. Peripheral proteins of postsynaptic membranes from Torpedo electric organ identified with monoclonal antibodies. J Cell Biol, 1984, 99:88~96
130 Mena EE, Foster AC, Fagg GE *et al*., Identification of synapse specific components: synaptic glycoproteins, protein, and transmitter binding sites. J Neurochem, 1981, 37:1557~1566
131 Palfrey HC, Rothlein JE, Greengard P. Calmodulin-dependent protein kinase and associated substrates in Torpedo electric organ. J Biol Chem, 1983, 258:9496~9503

132 Shields SM, Ingebritsen TS, Kelly PT. Identification of protein phosphatase 1 in synaptic junctions: dephosphorylation of endogenous calmodulin-dependent kinase II and synapse-enriched phosphoproteins. J Neurosci, 1985, 5:3414～3522
133 Sudhof TC. The synaptic vesicle cycle revisited. Neuron, 2000, 28: 317～320
134 Walikonis RS, Jensen ON, Mann M *et al*. Identification of proteins in the postsynaptic density fraction by mass spectrometry. J Neurosci, 2002, 20:4069～4080
135 Grant SG, Blackstock WP. Proteomics in neuroscience: from protein to network. J Neurosci, 2001, 21(21):8315～8318
136 Pawson T, Scott JD. Signaling through scaffold, anchoring, and adaptor proteins. Science, 1997, 278:2075～2082
137 Husi H, Ward MA, Choudhary JS *et al*. Proteomic analysis of NMDA receptor-adhesion protein signaling complexes. Nature America Inc, 2000, 3: 611～669
138 Hisatsune C, Umemori H, Mishina M *et al*. Phosphorylation-dependent interaction of the N-methyl-D-aspartate receptor epsilon 2 subunit with phosphatidylinositol 3-kinase. Gen Cells, 1999, 4:657～666
139 Craven SE, El-Husseini AE, Bredt DS. Synaptic targeting of the postsynaptic density protein PSD-95 mediated by lipid and protein motifs. Neuron, 1999, 22: 497～509
140 Mendelsohn AR, Brent R. Protein interaction methods-toward an endgame. Science, 1999, 284:1948～1950
141 Husi H, Grant SGN. Isolation of 2000kDa complexesof N-methyl-D-aspartate receptor and post synaptic density 95 from mouse brain. J Neurochem, 2001, 77:281～291
142 Grant SGN, Husi H. Proteomics of multiprotein complexes: answering fundamental questions in neuroscience. Trends Biotechnol, 2001, 19:49～52
143 Husi H, Grant SGN. Proteomics of the nervous system. Trends Neurosci, 2001, 24(5):259～266
144 Takashi O, Iku T, Kazuhiro K *et al*. Proteome analysis of rat hippocampal neurons by multiple large gel two-dimensional electrophoresis. Proteomics, 2002, 2:666～672
145 Gozal E, Gozal D, William MP *et al*. Proteomic analysis of CA1 and CA3 regions of rat hippocampus and differential susceptibility to intermittent hypoxia. J Neurochem, 2002, 83:331～345
146 Edgar PF, Douglas JE, Cooper GJ *et al*. Comparative proteome analysis of the hippocampus implicates chromosome 6q in schizophrenia. Mol Psychiat, 2000, 5(1):85～90
147 Wisniewaki HM, Rabe A. Discrepancy between alzhermer-type neuropathology and dementia in persons with down's syndrome. Ann NY Acad Sci, 1986, 477:247～260
148 Becker L, Mito T, Takashima S *et al*. Growth and development of the brain in down syndrome. Prog Clin Biol Res, 1991, 373:133～152
149 Cairns NJ. Neuropathology. J Neural Transm, 1999, 57:61～74
150 Nagy ZS. Mechanisms of neuronal death in down's syndrome. J Neural Transm, 1999, 57:233～245
151 Epstein CJ. Down syndrome(trisomy 21). In: Scriver CR *et al*. eds. The metabolic and molecular bases of inherited disease. New York: McGraw-Hill, 2001, 1223～1256
152 Cheon MS, Fountoulakis M, Dierssen M *et al*. Expression profiles of proteins in feta brain with down syndrome. J Neural Transm, 2001, 61:311～319
153 Longtine MS, DeMarin DJ, Balencik ML *et al*. The septins: roles in ctokinesis and other processes. Curr Opin Cell Biol, 1996, 8:106～119
154 Xie H, Surka M, Howard J *et al*. Characterization of the mammalian septin H5: distinct patterns of cytoskeletal and membrane association from other septin proteins. Cell Motil Cytoskeleton, 1999, 43:52～62
155 Kinoshita A, Noda M, Kinoshita M. Differential localization of septins in the mouse brain. J Comp Neurol, 2000, 428: 223～239
156 Yoo BC, Cairns N, Fountoulakis M *et al*. Synaptosomal protein, beta-soluble N-ethylmaleimide-sensitive factor attachment protein (beta-SNAP), gamma-SNAP and synaptotgmin I in brain of patients with down syndromen and Alzheimer's disease. Dement Geriatr Cogn Disord, 2001, 12:219～225
157 Markesbery WR. Oxidative stress hypothesis in Alzheimer's disease. Free Rad Biol Med, 1997, 23:134～147

158 Butterfield DA, Lauderback CM. Lipid peroxidation and protein oxidation in Alzheimer's disease brain: potential causes and consequences involving amyloid beta-peptide-associated free radical oxidative stress. Free Rad Biol Med, 2002, 32:1050～1060
159 Stadtman ER, Berlett BS. Reactive oxygen-mediated protein oxidation in aging and disease. Chem Res Toxicol, 1997, 10:485～494
160 Butterfield DA, Stadtman ER. Protein oxidation processes in aging brain. Adv Cell Aging Gerontol, 1997, 2:161～191
161 Butterfield DA, Drake J, Pocernich C *et al*. Evidence of oxidative damage in Alzheimer's disease brain: central role for amyloid beta-peptide. Trends Mol Med, 2001, 7:548～554
162 Smith CD, Carney JM, Starke-Reed PE *et al*. Protein oxidation in aging brain. Proc Natl Acad Sci USA, 1991, 88: 10540～10543
163 Hensley K, Hall N, Subramaniam R *et al*. Brain regional correspondence between Alzheimer's disease histopathology and biomarkers of protein oxidation. J Neurochem, 1995, 65:2146～2156
164 Castegna A, Aksenov M, Aksenova M *et al*. Proteomic identification of oxidatively modified proteins in Alzheimer's disease brain. Part Ⅰ: creatine kinase BB, glutamine synthase, ubiquitin carboxy-terminal hydrolase L-1, Free Rad Bioll Med, 2002, 562～571
165 Baulieu EE. Neurosteroids: a new function in the brain. Biol Cell, 1991, 71:3～10
166 Flood JF, Roberts E. Dehydroepiandrosterone sulfate improves memory in aging mice. Brain Res, 1988, 448:178～181
167 Weill-Engerer S, Jean-Philippe D, Sazdovitch V *et al*. Neurosteroid quantification in human brain regions: comparison between Alzheimer's and nondemented patients. J Clin Endocrinol Metab, 2002, 87(11):5138～5143
168 Schonberger SJ, Edgar PE, Kydd R *et al*. Proteomic analysis of the brain in Alzhemer's disease: molecular phenotype of a complex disease process. Proteomics, 2001, 1(12):1519～1528
169 Wolters ECH, Francot CMJE. Mental dysfunction in Parkinson's disease. Parkinsonism Relat Disord, 1998, 4:107～112
170 Langston JW, Ballard PAJR. Parkinson's disease in a chemist working with 1-methyl-4-phenyl-1, 2, 5, 6-tetrahydropyridine. N Engl J Med, 1983, 309(5):310
171 Kajita M, Niwa T, Takeda N *et al*. Presence of N-methyldopamine in parkinsonian and normal human brains. J Chromatogr, 1993, 613(1):1～8
172 Ikarashi Y, Maruyama Y. Identification of O-methyldopa in the ventricular fluid of patients with Parkinson's disease. Biol Mass Spectrom, 1992, 21(18):380～384
173 Zhou G, Shoji H, Yamada S *et al*. Decreased beta-phenylethylamine in CSF in Parkinson's disease. J Neurol Neurosurg Psychiatry, 1997, 63(6):754～758
174 Miura Y. Plasma beta-phenylethylamine in Parkinson's disease. Kurume Med J, 2000, 47(4):267～272
175 Wasinger VC, Cordwell SJ, Cerpa-Poljak A *et al*. Progress with gene-product mapping of the Mollicutes: Mycoplasma genitalium. Electrophoresis, 1995, 16(7):1090～1094
176 李伯良.功能蛋白质组学.生命的化学,1998,18(6):1～4
177 李林,吴家睿,李伯良.蛋白质组学的产生及其重要意义.生命科学,1999,30(6):533～539
178 李林. 蛋白质组学的进展.生物化学和生物物理学进展,2000,27(3):227～230
179 Anderson NL, Anderson NG. The human plasma proteome: history, character, and diagnostic prospects. Mol Cell Proteomics, 2002, 1: 845～867
180 Anderson NL, Anderson NG. A two-dimensional gel database of human plasma proteins. Electrophoresis, 1999, 12: 883～906
181 Hughes GJ *et al*. Plasma protein map: an update by microsequencing. Electrophoresis, 1992, 13: 707～714
182 Raida M, Schulz-Knappe P, Heine G *et al*. Liquid chromatography and electrospray mass spectrometric mapping of peptides from human plasma filtrate. J Am Soc Mass Spectrom, 1999, 10:45～54
183 Richter R *et al*. Composition of the peptide fraction in human blood plasma: database of circulating human peptides. J Chromatogr B Biomed Sci Appl, 1999, 726:25～35

184 Georgiou HM, Rice GE, Baker MS. Proteomic analysis of human plasma: failure of centrifugal ultrafiltration to remove albumin and other high molecular weight proteins. Proteomics, 2001, 1: 1503～1506

185 Adkins JN, Varnum SM, Auberry KJ *et al*. Toward a human blood serum proteome: analysis by multidimensional separation coupled with mass spectrometry. Mol Cell Proteomics, 2002, 1(12): 947～955

186 Musante L *et al*. Characterization of plasma factors that alter the permeability to albumin within isolated glomeruli. Proteomics, 2002, 2: 197～205

187 Petricoin EF *et al*. Use of proteomic patterns in serum to identify ovarian cancer. Lancet, 2002, 359: 572～577

188 Petricoin EF *et al*. Serum proteomic patterns for detection of prostate cancer. J Natl Cancer Inst, 2002, 94: 1576～1578

189 Bacarese-Hamilton T, Bistoni F, Crisanti A. Protein microarrays: from serodiagnosis to whole proteome scale analysis of the immune response against pathogenic microorganisms. Biotechniques, 2002, (Suppl): 24～29

190 Klade CS, VossT, Krystek E *et al*. Identification of tumor antigens in renal cell carcinoma by serological proteome analysis. Proteomics, 2001, 1(7): 890～898

191 Hanash S, Brichory F, Beer D. A proteomic approach to the identification of lung cancer markers. Dis Markers, 2002, 17(4): 295～300

192 Vytvytska O, Nage E, Bluggel M *et al*. Identification of vaccine candidate antigens of Staphylococcus aureus by serological proteome analysis. Proteomics, 2002, 2(5): 580～590

193 Lock RA, Coombs GW, McWilliams TM *et al*. Proteome analysis of highly immunoreactive proteins of Helicobacter pylori. Helicobacter, 2002, 7(3): 175～182

194 Patterson SD. Mass spectrometry and proteomics. Physiol Genomics, 2000, 2(2): 59～65

195 Davidsson P, Puchades M, Blennow K. Identification of synaptiç vesicle, pre-and postsynaptic proteins in human cerebrospinal fluid using liquid-phase isoelectric focusing. Electrophoresis, 1999, 20(3): 431～437

196 Rohlff C. Proteomics in molecular medicine: applications in central nervous systems disorders. Electrophoresis, 2000, 21(6): 1227～1234

197 Yuan X, Russell T, Wood G *et al*. Analysis of the human lumbar cerebrospinal fluid proteome. Electrophoresis, 2002, 23(7-8): 1185～1196

198 Choe LH, Green A, Knight RS *et al*. Apolipoprotein E and other cerebrospinal fluid proteins differentiate ante mortem variant Creutzfeldt-Jakob disease from ante mortem sporadic Creutzfeldt-Jakob disease. Electrophoresis, 2002, 23(14): 2242～2246

199 Burkhard PR., Rodrigo N, May D *et al*. Assessing cerebrospinal fluid rhinorrhea: a two-dimensional electrophoresis approach. Electrophoresis, 2001, 22(9): 1826～1833

200 Tseng WL, Lin YW, Chang HT. Improved separation of microhetero-geneities and isoforms of proteins by capillary electrophoresis using segmental filling with SDS and PEO in the background electrolyte. Anal Chem, 2002, 74(18): 4828～4834

201 Bergquist J, Palmblad M, Wetterhall M *et al*. Peptide mapping of proteins in human body fluids using electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry. Mass Spectrom Rev, 2002, 21(1): 2～15

202 Yuan X, Desiderio DM. Protein identification with Teflon as matrix-assisted laser desorption/ionization sample support. J Mass Spectrom, 2002, 37(5): 512～524

203 Cutillas PR, Norden AG, Cramer R *et al*. Detection and analysis of urinary peptides by on-line liquid chromatography and mass spectrometry: application to patients with renal Fanconi syndrome. Clin Sci (Lond), 2003 Jan 15, [epub ahead of print]

204 Kiernan UA, Tubbs KA, Nedelkov D *et al*. Comparative phenotypic analyses of human plasma and urinary retinol binding protein using mass spectrometric immunoassay. Biochem Biophys Res Commun, 2002, 297(2): 401～405

205 Sironi L, Tremoli E, Miller I *et al*. Acute-phase proteins before cerebral ischemia in stroke-prone rats: identification by proteomics. Stroke, 2001, 32(3): 753～760

206 Hampel DJ, Sansome C, Sha M *et al*. Toward proteomics in uroscopy: urinary protein profiles after radiocontrast

medium administration. J Am Soc Nephrol,2001,12(5):1026～1035
207 Vlahou A,Schellhammer PF,Mendrinos S. Development of a novel proteomic approach for the detection of transitional cell carcinoma of the bladder in urine. Am J Pathol,2001,158:1491～1500
208 Yao Y,Berg EA,Costello CE *et al*. Identification of protein components in human acquired enamel pellicle and whole saliva using novel proteomics approaches. J Biol Chem,2003,278 (7):5300～5308
209 Valenzuela JG. High-throughput approaches to study salivary proteins and genes from vectors of disease. Insect Biochem Mol Biol,2002,32(10):1199～1209
210 Ribeiro JM,Francischetti IM. Role of arthropod saliva in blood feeding: Sialome and Post-Sialome Perspectives. Annu Rev Entomol,2003,48: 73～88
211 Paweletz CP,Trock B,Pennanen M *et al*. Proteomic patterns of nipple aspirate fluids obtained by SELDI-TOF: potential for new biomarkers to aid in the diagnosis of breast cancer. Dis Markers,2001,17(4):301～307
212 Sauter ER,Zhu W,Fan XJ *et al*. Proteomic analysis of nipple aspirate fluid to detect biologic markers of breast cancer. Br J Cancer,2002,86(9):1440～1443
213 Rosty C,Christa L,Kuzdzal S *et al*. Identification of hepatocarcinoma-intestine-pancreas/pancreatitis-associated protein I as a biomarker for pancreatic ductal adenocarcinoma by protein biochip technology. Cancer Res,2002,62:1868～1875
214 Stark M,Jornvall H,Johansson J. Isolation and characterization of hydrophobic polypeptides in human bile. Eur J Biochem,1999,266(1):209～214
215 Sinz A,Bantscheff M,Mikkat S *et al*. Mass spectrometric proteome analyses of synovial fluids and plasmas from patients suffering from rheumatoid arthritis and comparison to reactive arthritis or osteoarthritis. Electrophoresis,2002,23(19):3445～3456
216 Arabidopsis Genome Initiative. Analysis of the genome sequence of the flowering plant Arabidopsis thaliana. Nature,2000,408:25～31
217 Picard P,Bourgoin-Greneche M,Zivy M. Potential of two-dimensional electrophoresis in routine identification of closely related durum wheat lines. Electrophoresis,1997,18:174～181
218 David JL,Zivy M,Cardin ML *et al*. Protein evolution in dynamically managed populations of wheat: adaptive responses to macro-environmental conditions. Theor Appl Genet,1997,95:932～941
219 Ramani S,Apte SK. Transient expression of multiple genes in salinity-stressed young seedlings of rice (Oryza sativa L.) cv. Bura Rata. Biochem Biophys Res Comm,1997,223:663～667
220 郭奕明,沈世华,荆玉祥等. Plant proteomics in the post-genomic era. 植物学报,2002,44(6):631～641
221 Tsugita A,Kawakami T,Uchiyama Y *et al*. Separation and characterization of rice proteins. Electrophoresis,1994,15:708～720
222 Chang WW,Huang L,Shen M *et al*. Patterns of protein synthesis and tolerance of anoxia in root tips of maize seedlings acclimated to a low-oxygen environment,and identification of proteins by mass spectrometry. Plant Physiol,2000,122:295～318
223 Shen S,Matsubae M,Takao T *et al*. A proteomic analysis of leaf sheaths from rice. J Biochem,2002,132(4):613～620
224 Komatsu S,Rakwal R,Li Z. Separation and characterization of proteins in rice (Oryza sativa) suspension cultured cells. Plant Cell Tiss Org Cult,1999,55:183～192
225 Bouchez D,Hoffe H. Functional genomics in plants. Plant Physiol,1998,118:725～732
226 van Wijk KJ. Proteomics of the chloroplast: experimentation and prediction. Trends Plant Sci,2000,5:420～425
227 Peltier JB,Friso G,Kalume DE *et al*. Proteomics of the chloroplast: systematic identification and targeting analysis of lumenal and peripheral thylakoid proteins. Plant Cell,2000,12:319～341
228 Yamaguchi K,Subramanian AR. The plastid ribosomal proteins. Identification of all the proteins in the 50S subunit of an organelle ribosome (chloroplast). J Biol Chem,2000,275:28466～28482
229 李学.药物蛋白质组学与药物研发.生理科学进展,2002,33(3):209～214

230 Dodds RA, James IE, Rieman D *et al*. Human osteoclast cathepsin K is processed intracellularly prior to attachment and bone resorption. J Bone Miner Res, 2001, 16(3): 478～486

231 Mullner S, Nermann T, Lottspeich F. Proteomics—a new way for drug target discovery. Arzneimittelforschung, 1998, 48(1): 93～95

232 Trevor CD. Signal transduction by the Wnt family of ligands. Biochem J, 1998, 329: 209～223

233 Anderson NL, Giere FA, Nance SL *et al*. Effects of toxic agents at the protein level: quatitative measurement of 213 mouse liver proteins following xenobiotic treatment. Fundam Appl Toxicol, 1987, 8(1): 39～50

234 Kanitz MH, Li EI. Investigation of ornnithine decarboxylase activity and two-dimensional electrophoretic protein profile following exposure of T24 bladder carcinoma cells to tumor promoter and carcinogen. Toxicol Methods, 1997, 7(1): 27～41

235 Aicher L, Wahl D, Arce A *et al*. New insights into cycloporine A nephrotoxicity by proteome analysis. Electrophoresis, 1998, 19(11): 1998～2003

236 Steiner S, Aicher L, Raymackers J *et al*. Decrease in kidney calbindin-D 28 kDa as a possible mechanism mediating cycloporine A. Biochem Pharmacol, 1997, 53: 723～731

后 记

科学出版社本来想组织《蛋白质化学研究技术与进展》一书第二版的写作，鉴于蛋白质组学的发展，我们将蛋白质化学的基本内容不变，增加了蛋白质组学的内容。但也延迟了一年时间。

十多年前我完全转入蛋白质这一领域，徐来根、吴高德、邵晓霞等和我并肩奋斗十多年，一批研究生在此得到良好的训练，同时也为该领域在我国的发展作出了贡献。其中有屠红旻、陈长征、车发云、俞利荣、曾嵘、宋金芳、刘韬、丁士健等，他们生长在改革开放的大好年代，研究环境有了极大的改善，是新人辈出的一代。

一个科学工作者所能做到的事情十分有限，我们无力与各种力量抗争，但可在允许的实验室里带一些学生耕耘，淡薄名利，陶冶人性。但受各种因素干扰还是在所难免。

此外，在本书撰写和出版过程中，科学出版社的谢灵玲、杨兆弘和莫结胜，中心的陈薇小姐和尤姝小姐给予全力的支持，在此特表感谢。

夏其昌

2003 年 9 月 24 日

2004.11.19
收到期 2004.12.8
来 源 [illegible]
书 价 52.50元
单据号 996442l
[illegible] 2004.12.2.

中科院植物所图书馆
S0003766

58.17421
468

000028123

(Q-1331.0101)

现代生物技术前沿丛书

书　名	书号　ISBN	定价
基因免疫的原理和方法	7-03-012588-6	38元
RNAi：基因沉默指南（影印版）	7-03-012685-8	65元
从基因到基因组 ——DNA技术概念和应用（影印版）	7-03-012477-4	46元
结构生物学与药学研究	7-03-011688-7	48元
生物芯片分析(影印版)	7-03-012247-X	80元
DNA芯片和基因表达（影印版）	7-03-012248-8	29元
计算分子生物学导论（翻译版）	7-03-011493-0	36元
植物生物技术导论（影印版）	7-03-012799-4	72元
蛋白质化学与蛋白质组学	7-03-012401-4	75元

ISBN 7-03-012401-4
9 787030 124012 >

生命科
联系电话　010
http://www
e-mail:info@life

ISBN 7-03-012401-
定 价：75.00 元

销售分类建议：生物技术/生物化学/分子生物学